

## THESIS / THÈSE

### MASTER EN BIOCHIMIE ET BIOLOGIE MOLÉCULAIRE ET CELLULAIRE

#### Rôle d'une modification post-transcriptionnelle de l' ARN dans le contrôle de la différenciation cellulaire (gamétogenèse)

Legraie, Sarah

*Award date:*  
2018

*Awarding institution:*  
Université de Namur

[Link to publication](#)

#### General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal ?

#### Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.



**Faculté des Sciences**

**RÔLE D'UNE MODIFICATION POST-TRANSCRIPTIONNELLE DE L'ARN DANS  
LE CONTRÔLE DE LA DIFFERENCIATION CELLULAIRE (GAMETOGENESE)**

**Mémoire présenté pour l'obtention  
du grade académique de master 120 en biochimie et biologie moléculaire et cellulaire**

SARAH LEGRAIE

Janvier 2018



## **Rôle d'une modification post-transcriptionnelle de l'ARN dans le contrôle de la différenciation cellulaire (gamétogenèse)**

LEGRAIE Sarah

### Résumé

Depuis plusieurs décennies, il est connu que l'ARN peut être modifié. Ces modifications, dites « post-transcriptionnelles », jouent des rôles structuraux et biologiques qui sont encore peu compris. Le champ d'étude qui vise à étudier et comprendre l'impact moléculaire et cellulaire de ces modifications, tout en développant des techniques de séquençage spécifiques, est appelé « épitranscriptomique ».

Ce mémoire s'intègre dans l'étude épitranscriptomique d'une modification correspondant à une forme réduite de l'uridine ; la dihydrouridine (D). La dihydrouridine est la modification de l'ARN la plus abondante au sein des ARN de transfert, et est synthétisée par des enzymes nommées dihydrouridine synthases.

Les expériences menées lors de ce mémoire s'articulent autour de deux objectifs. Premièrement, investiguer un effet biologique potentiel de la dihydrouridine en se basant sur un constat phénotypique préalable : un défaut de sporulation observé chez la levure de fission *Schizosaccharomyces pombe*, dans une souche délétée pour le gène codant pour l'enzyme dihydrouridine synthase 4 (Dus4). Un rôle potentiel de la dihydrouridine dans le phénomène de sporulation (gamétogénèse chez la levure) était alors envisageable. Deuxièmement, mettre au point une technique de séquençage, appelée « D-seq », permettant la détection spécifique de cette modification post-transcriptionnelle pour l'appliquer à deux autres espèces modèles : *Escherichia coli* et *Caenorhabditis elegans*. Le but de la réalisation du D-seq pour ces deux autres organismes est de pouvoir comparer phylogénétiquement la conservation des positions de la dihydrouridine au sein du transcriptome.

Les résultats des expériences et mises au point réalisés suggèrent que le défaut observé dans la sporulation chez *Schizosaccharomyces pombe* est indépendant de la protéine Dus4, et que le D-seq peut s'appliquer à deux autres espèces modèles à la biologie différente de *Schizosaccharomyces pombe*.



## Remerciements

La réalisation de ce mémoire n'aurait pas été rendue possible sans les présences et interventions de personnes que je tiens ici à remercier chaleureusement.

Premièrement, il me semble évident de remercier de manière la plus humble possible Monsieur Damien Hermand qui, il y a un an de cela, a accepté de m'ouvrir les portes de son laboratoire de Génétique Moléculaire. Je n'étais pourtant à l'époque qu'une étudiante inconnue sortant d'un bachelier de médecine vétérinaire qui venait bousculer l'attribution des sujets de mémoire, et ce de manière légèrement effrontée, je dois le reconnaître. Merci pour votre confiance, pour le partage de vos connaissances, et pour votre soutien tout au long de ces 10 mois de recherche, devrais-je dire de vie, avec leurs hauts et leurs bas. C'était pour moi un honneur de faire partie de votre équipe, et de vous côtoyer en tant que personne, tout simplement.

Si je m'étais permise à l'époque de remuer autant ciel et terre pour obtenir ce sujet de mémoire, cela était aussi dû à ma rencontre avec Olivier Finet, lors de la foire aux mémoires de novembre 2016. La génétique, « là où tout commence », ne m'avait jamais semblée aussi attrayante et intéressante qu'à l'écoute de tes mots passionnés. Merci de m'avoir pris sous ton aile, de m'avoir initiée au monde de la recherche scientifique, d'avoir toujours été à l'écoute et d'être la joie de vivre de ce laboratoire. Je me dois aussi de te remercier pour tes quelques remontrances, sans lesquelles je n'aurais pas pu essayer de m'améliorer et d'évoluer professionnellement parlant.

Je n'aurais pas pu non plus réaliser entièrement ce sujet de mémoire sans l'intervention de Fanélie Bauer. Merci pour les connaissances que tu m'as apportées sur ton organisme de prédilection. Même si je n'aurais finalement pas eu la peau de *Caenorhabditis elegans* lors de mes nombreux bleaching infructueux au cours de maintes week-end, c'était une expérience très enrichissante pour moi que de travailler pour la première fois avec un nouvel organisme modèle, et ce grâce à toi.

Merci aussi à Carlo Yague-Sanz, roi de la bio-informatique, qui grâce à ses connaissances, permet à l'étude épitranscriptomique de la dihydrouridine d'être interprétée et surtout comprise, chose très peu évidente à la base.

Merci de manière générale à tous les membres de l'équipe GéMo pour votre bonne humeur quotidienne, à Valérie pour ton éternelle pétillance, à Julie et Clément/Constant, maintenant devenus Docteurs.

Je tiens également à remercier les autres petits rats de laboratoire en apprentissage situés à quelques dizaines de mètres de mon bureau, et qui me manqueront à la fin de ces dix mois de recherche : Caroline, Audrey, Béa, Pauline, Nicolas. Merci pour votre amitié sincère, présence, soutien, et pour tous nos instants de détente. L'expérience « mémo » n'aurait pas été la même sans vous, et je suis très heureuse de pouvoir maintenant vous compter parmi mes amis.

Ensuite, un grand merci à mes parents qui, même si les simples mots « modifications post-transcriptionnelles » ou « génétique moléculaire » les font frémir d'interrogations, se sont toujours montrés présents. Et ce tout au long de mon parcours universitaire. Merci de m'avoir soutenue dans mes choix passés, présents et futurs. Merci de me permettre de me construire l'avenir professionnel auquel j'aspire.

Merci aussi à ma famille et autres amis du quotidien, qui m'apportent confiance, et ce tout autant que les ronronnements apaisants de mon chat Kiwi sur les articles scientifiques que j'avais à décortiquer tout au long de ces derniers mois.

Merci au Jury de ce mémoire : O. De Backer, D. Coupeau, R.Hallez, P. Renard, de prendre temps et considération pour valoriser mon projet de recherche réalisé au sein du laboratoire de Génétique Moléculaire de l'UNamur.

Enfin, merci à l'Université de Namur qui m'a permis de réaliser ce parcours universitaire particulier. En aucun cas je ne reviendrais sur mon choix d'avoir finalement entrepris un master en Biologie au sein de notre belle capitale wallonne.

Sasatouille.



**Rôle d'une modification post-transcriptionnelle de l'ARN dans le contrôle de la différenciation cellulaire (gamétogenèse) :**

**La dihydrouridine**

**Implication de la dihydrouridine synthase Dus4 dans la sporulation chez *Schizosaccharomyces pombe***

**&**

**Préparation d'une analyse épitranscriptomique comparative de la dihydrouridine chez *Schizosaccharomyces pombe*, *Escherichia coli* et *Caenorhabditis elegans***

## Table des matières

I) Introduction	8
I.1 Généralités sur les modifications post-transcriptionnelles	8
I.2 La dihydrouridine et les enzymes dihydrouridine synthase (Dus)	9
I.3 Constats préalables au laboratoire GéMo (Génétique Moléculaire, UNamur) et émergence du projet	11
I.3.1 Une espèce modèle importante en génétique : <i>Schizosaccharomyces pombe</i>	11
I.3.2 Constats préalables clés	11
I.3.3 La sporulation chez <i>Schizosaccharomyces pombe</i>	12
I.3.4 Le gène <i>spo6</i>	13
I.4 Méthodes de détection des modifications post-transcriptionnelles	14
I.4.1 Historique	14
I.4.2 Généralités	14
I.4.3 Exemples	15
• N <sup>6</sup> -méthyladénosine (m <sup>6</sup> A)	
• Pseudouridine (ψ)	
I.4.4 Difficultés	16
I.5 Expérience préalable de détection de la dihydrouridine chez <i>Schizosaccharomyces pombe</i> ou « D-seq » et suite du projet	17
I.6 Choix de deux autres espèces modèles	18
I.6.1 <i>Escherichia coli</i>	
I.6.2 <i>Caenorhabditis elegans</i>	
II) Objectifs	21
III) Matériels et Méthodes	22
III.1 Implication de la dihydrouridine synthase Dus4 dans la sporulation chez <i>Schizosaccharomyces pombe</i>	22
III.2 Préparation d'une analyse épitranscriptomique comparative de la dihydrouridine chez <i>Schizosaccharomyces pombe</i> , <i>Escherichia coli</i> et <i>Caenorhabditis elegans</i>	25
IV) Résultats	27
IV.1 Implication de la dihydrouridine synthase Dus4 dans la sporulation de <i>Schizosaccharomyces pombe</i>	27
IV.1.1 Observation d'un défaut de sporulation dans la souche $\Delta dus4$ corrélé avec une diminution de l'expression du gène <i>spo6</i>	27



IV.1.2 Etude de l'implication de l'activité catalytique de la protéine Dus4 dans le défaut de sporulation observé et corrélé avec une diminution de l'expression du gène <i>spo6</i> dans la souche $\Delta dus4$	28
IV.1.3 Etude de l'implication de la protéine Dus4 dans le défaut de sporulation observé dans la souche $\Delta dus4$	30
IV.1.4 Etude de l'implication du transcrit <i>dus4</i> dans le défaut de sporulation observé dans la souche $\Delta dus4$	30
IV.1.5 Etude de l'implication de la cassette de résistance utilisée dans la souche $\Delta dus4$ dans le défaut de sporulation observé	32
IV.1.6 Etude de l'implication de l'environnement génomique dans le défaut de sporulation observé dans la souche $\Delta dus4$	33
IV.1.7 Etude de la restauration de la sporulation « wild type » observée dans la souche $\Delta dus4 \Delta dus2$	33
IV.2 Préparation d'une analyse épitranscriptomique comparative de la dihydrouridine chez <i>Schizosaccharomyces pombe</i> , <i>Escherichia coli</i> et <i>Caenorhabditis elegans</i>	34
IV.2.1 Obtention, contrôles, culture et récolte des souches d' <i>Escherichia coli</i> et <i>Caenorhabditis elegans</i>	35
<ul style="list-style-type: none"> <li>• <i>Escherichia coli</i></li> <li>• <i>Caenorhabditis elegans</i></li> </ul>	
IV.2.2 Tests d'extraction d'ARN et analyses réalisés sur les souches d' <i>Escherichia coli</i> et <i>Caenorhabditis elegans</i> récoltées	36
IV.2.3 « Rhodamine labelling » sur les ARN extraits d' <i>Escherichia coli</i> et <i>Caenorhabditis elegans</i>	36
IV.2.4 Etapes du « D-seq »	37
IV.2.5 Préparation des librairies à envoyer au séquençage	38
V. Discussion	40
V.1 Implication de la dihydrouridine synthase Dus4 dans la sporulation chez <i>Schizosaccharomyces pombe</i>	40
V.2 Préparation d'une analyse épitranscriptomique comparative de la dihydrouridine chez <i>Schizosaccharomyces pombe</i> , <i>Escherichia coli</i> et <i>Caenorhabditis elegans</i>	44
VI. Conclusion et perspectives	47
VII. Annexes	48
Table 1 : Souches de <i>Schizosaccharomyces Pombe</i> utilisées dans ce mémoire	
Table 2 : Primers utilisés dans ce mémoire	
VIII. Bibliographie	51



## I) Introduction



## D) Introduction

### I.1 Généralités sur les modifications post-transcriptionnelles

C'est en 1957 que la première modification post-transcriptionnelle de l'ARN a été identifiée, suite à une analyse d'un extrait d'ARN par chromatographie à échange d'ions<sup>1</sup>. En s'attendant à mettre en évidence les quatre nucléotides composant l'ARN (adénosine / cytidine / guanosine / uridine) une cinquième entité surnommée le « cinquième nucléotide » fut découverte. Plus tard, ce cinquième nucléotide a été identifié comme étant la pseudouridine<sup>2</sup>. La pseudouridine est une modification de l'ARN parmi bien d'autres et ensemble, elles constituent ce qui est maintenant appelé « l'épitranscriptome ». Tout ARN peut être concerné (ARN messager/ARNm, ARN ribosomique/ARNr, ARN de transfert/ARNt, long ARN non codant/ARNlnc, petit ARN nucléaire/ARNsn, petit ARN nucléolaire/ARNsno<sup>1</sup>). L'étude des modifications post-transcriptionnelles s'est initialement peu développée, étant donné le challenge technique que représente l'analyse et la caractérisation de nouvelles modifications<sup>3</sup>, et leur faible prévalence comparativement à la quantité de nucléotides standards<sup>4</sup>.

Depuis, grâce au développement croissant des méthodes d'étude, plus d'une centaine de modifications chimiques de l'ARN ont été répertoriées. Toutes ces modifications post-transcriptionnelles sont référencées dans la base de données MODOMICS<sup>5</sup>. A l'heure actuelle, cette base de données contient des informations structurales et biologiques sur 170 modifications, et ce nombre augmente continuellement<sup>5</sup>. Ces modifications peuvent par exemple être des méthylations, réductions, oxydations ou échanges de groupes et peuvent avoir lieu au niveau de la base, du 2'-hydroxyl du ribose, ou à ces deux endroits simultanément<sup>3</sup>. Les six modifications chimiques les mieux caractérisées de nos jours dans le monde de l'épitranscriptomique sont la N<sup>6</sup>-méthyladénosine (m<sup>6</sup>A), l'inosine (I), la 5-méthylcytosine (m<sup>5</sup>C), la pseudouridine ( $\psi$ ), la 5-hydroxyméthylcytosine (hm<sup>5</sup>C) et la N<sup>1</sup>-méthyladénosine (m<sup>1</sup>A)<sup>2</sup>. Il est maintenant établi que chaque acide nucléique dans la cellule puisse subir une forme de modifications post-transcriptionnelle<sup>3</sup>.

Ces modifications post-transcriptionnelles apparaissent dans les trois domaines de la vie (eucaryotes, archées, bactéries)<sup>3</sup>. Bien que leurs localisations, abondances et distributions dépendent fortement de la molécule d'ARN concernée, de l'organisme et de l'organite<sup>5</sup>, certaines de leurs positions s'avèrent être parfois évolutivement très conservées<sup>3</sup>. C'est par exemple le cas d'une pseudouridine présente en position 55 des ARNt<sup>3</sup>, du m<sup>6</sup>A particulièrement situé au pourtour du codon STOP<sup>4</sup>, du m<sup>1</sup>A au pourtour du codon start AUG<sup>6</sup> ou d'uridines modifiées situées en première position de l'anticodon des ARNt<sup>3</sup>. Cette conservation interspécifique suggère que ces modifications constituent un trait adaptatif biologiquement relevant et sélectionnable au cours de l'évolution.

Toute modification post-transcriptionnelle est catalysée par l'action d'enzymes. D'ailleurs, chez la levure de bourgeonnement *Saccharomyces cerevisiae*, approximativement 1% du génome code pour des protéines impliquées dans la biosynthèse de modifications post-transcriptionnelles<sup>3</sup>. La mise en place d'une modification sur une base nucléotidique se fait par l'action d'une enzyme spécifique, qui peut parfois appartenir à un complexe de plusieurs protéines, appelé « writer ». Au cours de la dernière décennie, plus de 40 gènes codant pour des writers ont pu être identifiés<sup>7</sup>, et ce grâce à l'utilisation de méthodes de génétique inverse et de la spectrométrie de masse. D'autres enzymes, appelées « eraser », peuvent retirer ces modifications post-transcriptionnelles, qui s'avèrent alors être réversibles. Cette dynamique supporte aussi leur relevance biologique.



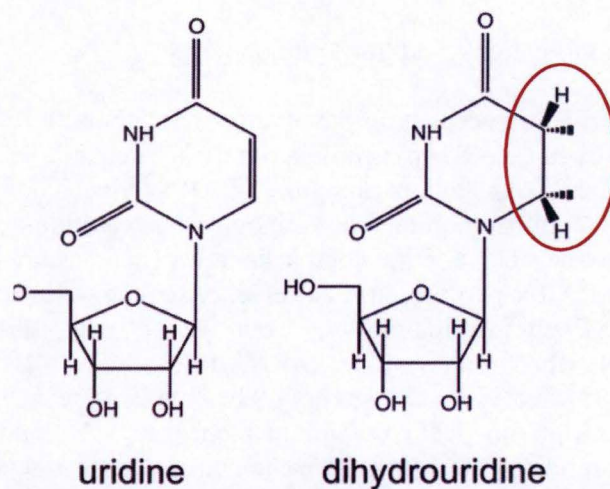


Figure 1. Structure chimique de l'uridine et de la dihydrouridine. Le lien C5-C6 de l'uridine est réduit pour former la dihydrouridine <sup>8</sup>.

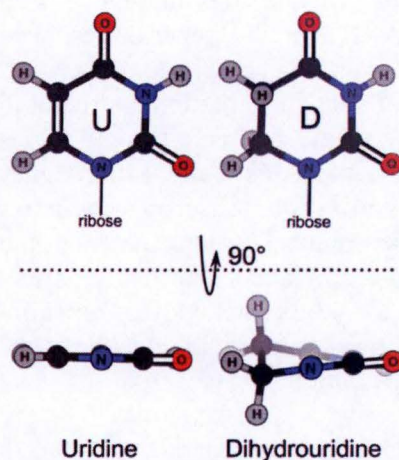


Figure 2. Modification de la conformation de l'uridine en dihydrouridine. La base formée est non plane <sup>9</sup>.

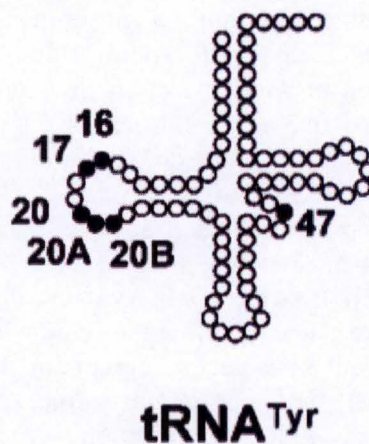


Figure 3. Profil des positions de la dihydrouridine d'un ARN de transfert (tRNA<sup>Tyr</sup>) obtenu dans l'étude de Xing *et al.* <sup>10</sup> La dihydrouridine a été identifiée en position 16, 17, 20, 20A, 20B et 47 <sup>10</sup>.



Enfin, dans les cas de la modification m<sup>6</sup>A (voir plus loin), des enzymes appelées « reader » ont aussi été identifiées. Elles constituent les médiateurs des effets de la modifications post-transcriptionnelle notamment par le biais de protéines spécifiques de liaison au méthyl<sup>6</sup>. Un seul site au sein de l'ARNm peut donc alors être modifié, ou restauré tout au long de la vie de l'ARN, ce qui en permet une régulation fine au sein de la cellule<sup>8</sup>.

A l'inverse de l'épigénome qui concerne les modifications de la chromatine, l'épitranscriptome confère des modifications dont l'effet ne persiste pas dans la cellule étant donné le turnover continu subi par les ARN<sup>7</sup> (la plupart des ARNm ont un temps de vie moyen de quelques heures<sup>8</sup>). De plus, l'ARN peut être présent en plusieurs copies dans la cellule, ce qui pose la question de la proportion des molécules modifiées. A contrario de l'ADN, qui lui peut voir ses sites modifiés uniquement une ou deux fois (voir aucune)<sup>7</sup>.

Basiquement, les modifications post-transcriptionnelles peuvent impacter la cellule de deux manières : biologiquement et structurellement. Biologiquement, les modifications de l'ARN peuvent impacter tous les aspects de la vie de l'ARN tels que l'épissage, la localisation de l'ARN, sa traduction et sa dégradation<sup>6</sup>. Structurellement, les modifications post-transcriptionnelles permettent à l'ARN d'adopter une conformation particulière. Cela est observé de manière importante au sein des ARNr et des ARNt.

Il est important de noter que toutes ces modifications post-transcriptionnelles font partie d'un ensemble de signaux de régulation<sup>3</sup> (par exemple via les protéines d'interaction avec l'ARN<sup>9</sup>) et agissent de manière coordonnée pour connecter signaux environnementaux et réponses cellulaires. Ces signaux peuvent être physiologiques (croissance, différenciation cellulaire), pathologiques (cancers) ou extérieurs (stress, choc thermique). De plus, cette régulation apparaît cruciale pour la cellule. Par exemple, il a pu être mis en évidence que la présence de modifications post-transcriptionnelles aberrantes provoquait un effet dévastateur lors du stade de développement précoce chez l'humain et la souris, de l'inflammation, de la neurodégénération et de maladies cancéreuses<sup>7</sup>.

Il est donc clair que l'étude du domaine de l'épitranscriptomique n'en est encore qu'à ses débuts et que les modifications post-transcriptionnelles de l'ARN ont encore beaucoup à nous apprendre dans le monde de la génétique, et ce au fur et à mesure des études et découvertes futures.

## I.2 La dihydrouridine et les enzymes dihydrouridine synthase (Dus)

Ce mémoire se concentre sur la dihydrouridine (D ; 5,6-dihydro-uridine). La dihydrouridine est la modification post-transcriptionnelle la plus représentée au sein des ARN de transfert<sup>10</sup>. Elle l'est particulièrement au sein des ARNt des eubactéries et des eucaryotes, et un peu moins abondamment au sein des ARNt des archéobactéries<sup>10</sup>. Dans les ARNt, environ 10% des nucléosides sont modifiés<sup>11</sup>, 105 modifications ayant été répertoriées jusqu'à présent<sup>12</sup>. Ces modifications affectent la chimie locale, conformation et flexibilité de l'ARNt<sup>13</sup>.

La dihydrouridine est formée à partir d'uridine, suite à la réduction de son double lien C5-C6 (**Fig. 1**). Cela provoque la formation d'une base non plane qui empêche dès lors un empilement normal des bases azotées (**Fig. 2**)<sup>14</sup>. De manière très conservée, la dihydrouridine est majoritairement présente au sein du D-loop (Dihydrouridine-loop) des ARN de transfert à des positions spécifiques (majoritairement en positions 16, 17, 20, 20a et 20b) et en position 47 du bras variable de l'ARNt<sup>15</sup> (**Fig. 3**).



Des analyses NMR (nuclear magnetic resonance) montrent que la dihydrouridine déstabilise la structure de l'ARNt en favorisant la conformation du ribose C2'-endo plutôt que la conformation C3'-endo, ce qui est thermodynamiquement favorisé. Ce changement de conformation augmente probablement la flexibilité et la dynamique des régions de l'ARN qui jouent un rôle dans les interactions 3D. De ce fait, la dihydrouridine se retrouve principalement au niveau des boucles simple brin d'ARN et au niveau de zones de la chaîne nucléotidique de l'ARN où une forte tension est exercée<sup>16</sup>. Étant donné que la dihydrouridine augmente la flexibilité des ARN de transfert, il est proposé qu'elle serait d'autant plus importante et abondante lors du folding des ARNt des bactéries psychrophiles et mésophiles<sup>17</sup>, en conditions de faibles températures<sup>14</sup>. Il a aussi été proposé que la dihydrouridine jouerait un rôle dans la fidélité de la traduction en augmentant l'interaction avec les partenaires de l'aminoacyl ARNt synthetase<sup>13</sup>.

La formation de la dihydrouridine est catalysée par l'action de « writers » appelés dihydrouridine synthases (Dus). Différentes familles de dihydrouridine synthases existent, chaque Dus étant responsable de la formation d'une dihydrouridine à une position particulière de l'ARN de transfert. Par exemple, quatre dihydrouridine synthases ont été référencées chez *Saccharomyces cerevisiae* : Dus1, Dus2, Dus3, Dus4 et trois chez *Escherichia coli* : DusA, DusB, DusC<sup>11</sup>. Des analyses biochimiques ont montré que les Dus étaient des enzymes flavine-dépendantes, nécessitant du NADPH (Nicotinamide adénine dinucléotide phosphate) ou du NADH (Nicotinamide adénine dinucléotide) pour l'activité catalytique<sup>18</sup>. Deux autres études se sont aussi plus particulièrement penchées sur la détermination des composants moléculaires de l'enzyme étant responsables de son activité enzymatique (notamment la cystéine en position 108 et la lysine en position 149)<sup>19, 20</sup>. Les Dus constituent une famille conservée d'enzymes encodées par une même famille de gènes. Ces gènes sont retrouvés dans tous les organismes vivants dont le génome a été séquencé<sup>16</sup>. À l'heure actuelle, des cristallographies X-ray sont disponibles pour trois enzymes Dus<sup>21,11,19</sup>, ce qui permet de visualiser et de comprendre l'interaction physique de ces enzymes avec leur substrat.

Concernant le rôle biologique de la dihydrouridine, peu de choses sont connues. Une étude de 2005<sup>22</sup> a rapporté qu'une surexpression de la dihydrouridine synthase 2 humaine était corrélée à la carcinogénèse pulmonaire. Suite à cela, une équipe de recherche a modélisé en 2015 la structure 3D de cette enzyme afin de pouvoir potentiellement aider au développement de nouvelles approches thérapeutiques dans le combat contre le cancer du poumon<sup>13</sup>.

Jusqu'à présent, peu d'études se sont donc consacrées à la dihydrouridine et à ses writers (Dus). Cependant, étant donné leur conservation importante au sein des espèces, investiguer plus en profondeur leur effets potentiels sur l'ARN et sur la biologie de la cellule apporterait potentiellement des informations pertinentes dans le domaine de l'épitranscriptomique. D'où l'intérêt de ce mémoire.



### I.3 Constats préalables au laboratoire GéMo (Génétique Moléculaire, UNamur) et émergence du projet

#### I.3.1 Une espèce modèle importante en génétique : *Schizosaccharomyces pombe*

Une grande partie de ce mémoire réalisée dans le laboratoire de Génétique Moléculaire de l'Université de Namur se base sur l'utilisation de l'espèce modèle *Schizosaccharomyces pombe*. Il s'agit d'une levure utilisée particulièrement en génétique depuis les années 1940. C'est une très lointaine cousine de *Saccharomyces cerevisiae*, la levure de boulangerie. *S. pombe* est une levure ascomycète appartenant au règne des mycètes. C'est un eucaryote unicellulaire capable de fermenter les sucres. Elle est aussi appelée levure de fission (contrairement à *Saccharomyces cerevisiae* surnommée la levure de bourgeonnement), en référence à son cycle de vie. En tant qu'eucaryote, la levure de fission peut être utilisée pour étudier des phénomènes cellulaires qui sont conservés entre la levure et l'humain mais absents des bactéries, comme la biogénèse des organites, l'organisation du cytosquelette, certaines spécificités de la transcription, de la réplication de l'ADN et de la traduction. Le modèle *S. pombe* est célèbre pour l'identification des régulateurs du cycle cellulaire et de la formation de l'hétérochromatine <sup>23</sup>.

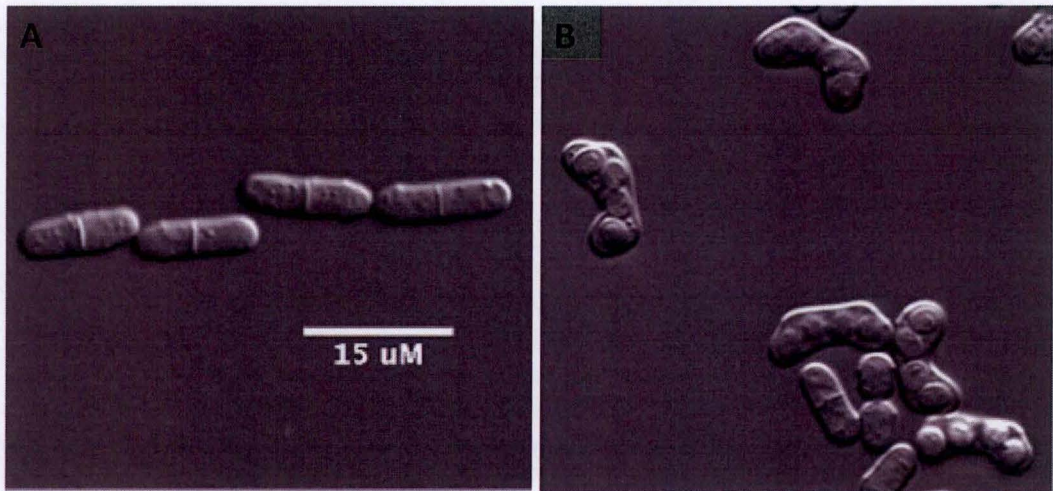
*S. pombe* s'avère être très intéressante dans le domaine de la recherche en génétique moléculaire pour plusieurs raisons. Les souches utilisées en laboratoire peuvent être haploïdes ou diploïdes, ce qui est utile pour mettre en évidence la récessivité des allèles. Il est aussi possible de transformer des plasmides auto-réplicatifs dans ces levures. De plus, *S. pombe* possède des mécanismes de recombinaison homologue hautement efficaces, ce qui permet la construction relativement aisée de nouvelles souches dont le génome est modifié <sup>24</sup>. La reproduction sexuée permet de facilement combiner des mutations. Enfin, un autre avantage dans le monde de la recherche sur *S. pombe* est que toutes les souches répertoriées dérivent de trois souches initiales (968 *h*<sup>90</sup>, 972 *h*<sup>-</sup>, 975 *h*<sup>+</sup>), ce qui fait que les souches utilisées en laboratoire se trouvent être toutes isogéniques <sup>24</sup>.

La taille du génome de *S. pombe* est d'approximativement 13,8 Mb, il est compris au sein de 3 larges chromosomes sur lesquels 5054 gènes codant pour des protéines ont été référencés. A l'heure actuelle, 2154 de ces protéines possèdent un rôle biologique publié et 338 de ces protéines s'avèrent être conservées parmi les vertébrés <sup>24</sup>. En ce qui concerne l'ARN, en plus des 307 ARNr identifiés, des ARNt, des ARNsn et des ARNsno, 1538 ARN non codants sont actuellement annotés dans le génome de *S. pombe*. Et tous ces gènes sont référencés dans une base de données universelle : [www.pombase.org](http://www.pombase.org) <sup>24</sup>.

#### I.3.2 Constats préalables clés

Au cours d'un projet de recherche mené au laboratoire, une souche délétée pour le gène codant pour l'enzyme dihydrouridine synthase 4 (souche  $\Delta$ *dus4*) et une souche délétée en plus pour le gène codant pour l'enzyme dihydrouridine synthase 2 (souche  $\Delta$ *dus4*  $\Delta$ *dus2*) ont été construites. Dans le cadre d'un autre projet de recherche, le transcriptome de ces variants a été séquencé (RNAseq) afin de détecter des variations d'expression comparativement à une souche non modifiée (wild type, WT). Il s'est avéré après analyse des données obtenues par RNAseq qu'un cluster de gènes était sous-exprimé dans la souche  $\Delta$ *dus4* et que ce même cluster de gènes était de nouveau exprimé à un niveau comparable à celui de la souche wild type dans la souche  $\Delta$ *dus4*  $\Delta$ *dus2*.





**Figure 4. Cellules de *Schizosaccharomyces pombe*.** A) Cellules haploïdes en fin de croissance mitotique et contenant un septum. B) Asques zygotiques formés suite à la conjugaison de deux cellules haploïdes, chacun contenant 4 spores <sup>24</sup>.



Parmi ce cluster, un gène nécessaire à la sporulation de la levure a été mis en évidence : le gène *spo6*. Pour vérifier cet effet, les souches  $\Delta dus4$  et  $\Delta dus4 \Delta dus2$  ont été observées par microscopie. Il s'est avéré que la souche  $\Delta dus4$  présentait en effet un défaut de sporulation et que la souche  $\Delta dus4 \Delta dus2$  sporulait de façon similaire à une souche wild type. Cela suggérait que la délétion du gène *dus2* supprimait le phénotype du mutant  $\Delta dus4$ , et posait la question du lien entre dihydrouridine et sporulation.

Lors de l'étude d'une modification post-transcriptionnelle, deux éléments sont importants : comprendre son effet biologique et établir sa distribution au sein du transcriptome. Le but de la première partie de ce mémoire était d'étudier le rôle biologique potentiel de la dihydrouridine suite à l'observation du défaut de sporulation dans la souche  $\Delta dus4$ , ceci en approfondissant et en quantifiant les analyses préliminaires. Ensuite, en construisant et analysant tout un nombre d'autres souches afin de comprendre quelles parties de la protéine ou de la construction pouvaient être responsables du phénotype observé.

Afin de permettre la compréhension du défaut de sporulation observé dans la levure de fission *S. pombe*, des données théoriques sont ici apportées concernant le cycle de vie de la levure et les informations actuellement disponibles sur le gène de sporulation *spo6*.

### 1.3.3 La sporulation chez *Schizosaccharomyces pombe*

En fonction des conditions environnementales, *S. pombe* poursuit un cycle asexué (forme végétative), ou entre en gamétogenèse (phase sexuée). En condition végétative, dans un milieu riche en nutriments, la levure de fission prolifère par divisions mitotiques.

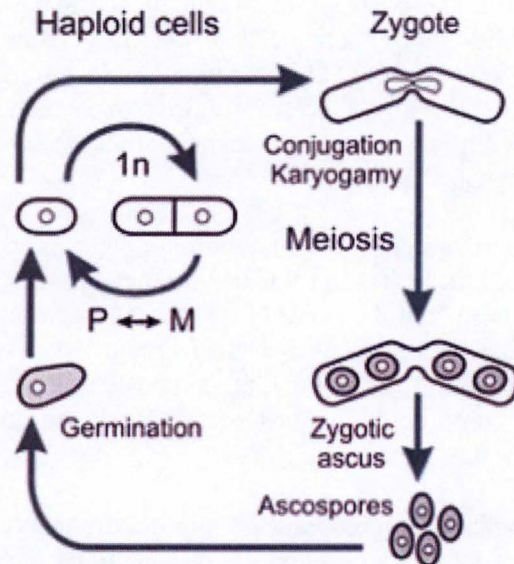
*S. pombe* est de forme cylindrique aux extrémités arrondies et existe majoritairement sous forme haploïde (la forme diploïde existe mais est plus rare). Les cellules post-mitotiques mesurent approximativement 8  $\mu\text{m}$  de long et atteignent environ 15  $\mu\text{m}$  de long en fin de croissance, moment où elles entrent en mitose<sup>24</sup>. La longueur d'une cellule wild type peut donc être un indicateur de sa progression au sein du cycle cellulaire<sup>25</sup>.

Les deux cellules filles formées après mitose sont de tailles équivalentes, séparées par un septum, et subiront un mode de division symétrique, leur génome étant équivalent à celui de leur cellule parente<sup>24</sup> (**Fig. 4**). La phase G<sub>1</sub> a lieu entre la mitose et la séparation des deux cellules filles (cytokinèse). La phase S se déroule au moment où le septum est formé et les cellules filles séparées sont donc déjà en phase G<sub>2</sub>, contenant des chromosomes répliqués. Bien que le cycle soit haploïde, les cellules sont donc fonctionnellement diploïdes<sup>24</sup>.

Quand les cellules sont soumises à une carence en nutriments (c'est-à-dire un stress), et en l'absence de partenaire, elles sortent du cycle mitotique et entrent en phase stationnaire. Si un partenaire sexuel est présent, les cellules haploïdes initient différenciation sexuelle<sup>24</sup>.

Lors de la différenciation sexuelle, la cellule arrête le cycle cellulaire en phase G<sub>1</sub> et conjugue avec une cellule du « mating type » opposé<sup>26</sup>, grâce à des phéromones de conjugaison différentes émises en fonction du mating type. Les souches naturelles de *S. pombe* sont dites « homothallique *h90* ». Génétiquement, elles possèdent les cassettes spécifiant les deux types sexuels (*h*<sup>-</sup>, *h*<sup>+</sup>) et une d'entre elles est aléatoirement exprimée tandis que l'autre est transcriptionnellement réprimée sous forme d'hétérochromatine. De plus, ces souches sont capables de changer de type sexuel (« mating type switching »)<sup>24</sup>.





**Figure 5. Cycle de vie de *Schizosaccharomyces pombe*.** Sans stress environnemental, les cellules végétatives haploïdes subissent des divisions mitotiques (1n vers 1n). En conditions de carences, les cellules haploïdes expriment un gène du mating type : une cellule P (plus) sécrète un facteur P, une cellule M (moins) sécrète un facteur M. La cellule P produit un récepteur au facteur M tandis que la cellule M produit un récepteur au facteur P. Ce processus est appelé différenciation sexuelle. Ensuite, la conjugaison a lieu entre une cellule P et une cellule M. Le zygote formé (diploïde 2n) subit la méiose I et II et résulte en la formation de 4 spores empaquetées dans un asque zygotique. Les spores peuvent être définies comme des cellules haploïdes résistantes. Quand l'environnement devient favorable, les spores sont libérées de l'asque et germent pour devenir des cellules végétatives haploïdes. Figure adaptée du site web de O.Nielsen.



En laboratoire, il existe des souches qui n'expriment soit que le mating type *h+* ou soit le mating type *h-*. Ces souches sont dites « hétérothalliques », chacune possédant un récepteur à la phéromone de conjugaison émise par la souche du mating type opposé. Les souches haploïdes *h90* peuvent conjuguer avec n'importe quelle autre souche avant de subir la méiose et les souches *h+* conjuguent donc avec les souches *h-* et inversement. Suite à la conjugaison de deux cellules de mating type opposé, un zygote diploïde est formé et peut entrer en méiose.

Les différentes étapes de la méiose chez *S. pombe* sont similaires à ce qui est observé chez les eucaryotes multi-cellulaires, ce qui fait d'ailleurs de la levure de fission un des meilleurs organismes modèles pour étudier les mécanismes régulateurs de la méiose<sup>26</sup>. Dans le zygote formé après conjugaison, la synthèse d'ADN préméiotique est suivie par deux étapes de division nucléaire (méiose I et méiose II). Les quatre nuclei produits sont ensuite encapsulés par une paroi pour former des ascospores (Fig. 4)<sup>26</sup>. Les spores sont ensuite libérées, et germent quand les conditions redeviennent favorables pour rentrer dans un cycle de division végétatif. Le processus de gamétogenèse est caractérisé par des recombinaisons inter- et intra-chromosomiques. Le cycle de la levure de fission est schématisé figure 5 (Fig. 5). Deux exigences doivent donc être rencontrées pour permettre d'initier la méiose dans des cellules *S. pombe* wild type : une carence en nutriments, plus particulièrement en azote, et une ploïdie correcte<sup>26</sup>.

Chez *S. pombe*, le facteur de transcription principalement responsable du passage de la prolifération cellulaire à la différenciation sexuelle est encodé par le gène *ste11*<sup>27</sup>. L'expression de *ste11* est sous le contrôle de plusieurs voies de signalisations qui répondent à des facteurs externes comme la quantité de nutriments, les stress et les phéromones de conjugaison (voies cAMP – cyclic adenosine monophosphate, TORC1 – target of rapamycin complex 1, MAPK mitogen – activated protein kinase)<sup>27</sup>.

Dans ce mémoire, la technique de croisement a été utilisée pour la constructions de plusieurs souches. Le croisement de deux cellules haploïdes consiste en leur conjugaison sur un milieu pauvre en azote qui aboutira à un zygote diploïde. Suite à la méiose et au brassage du matériel génétique, 4 spores haploïdes seront formées. Grâce à ce brassage, de nouvelles combinaisons alléliques sont générées en suivant les lois de la génétique mendélienne. Ces combinaisons sont ensuite identifiées par sélection (voir matériels et méthodes).

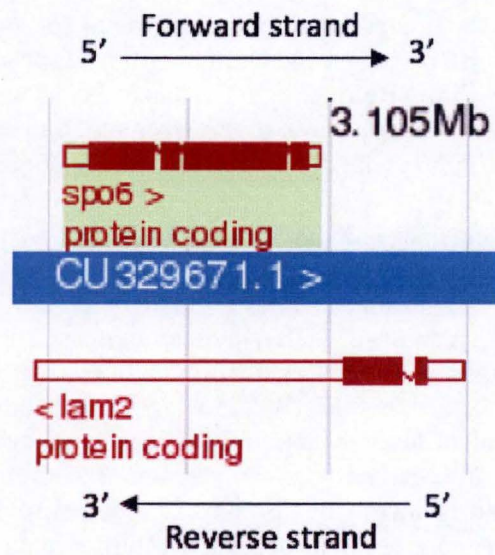
Ici, le défaut de sporulation observé quand le gène *dus4* est absent se manifeste par des spores dont la morphologie et le nombre sont aberrants.

#### 1.3.4 Le gène *spo6*

Pour rappel, dans la souche  $\Delta$ *dus4* observée et présentant un défaut de sporulation, il est aussi apparu lors de l'analyse en RNAseq que le gène de sporulation *spo6* était sous-exprimé.

Une étude<sup>28</sup> chez *S. pombe* a mis en évidence que *spo6* était un gène requis pour la sporulation et dont la transcription était régulée par le facteur de transcription Mei4. Une autre étude a identifié 20 loci comme ayant une importance dans la sporulation chez la levure de fission (*spo1-spo20*)<sup>29</sup>. De plus, dans cette étude<sup>28</sup>, un mutant délété pour le gène *spo6* présentait aussi un défaut lors du phénomène de sporulation. En analysant plus précisément le cycle de ce mutant, il s'est avéré que *spo6* était nécessaire lors de la seconde division méiotique, la majorité des mutants homozygotes diploïdes pour *spo6* s'arrêtant au stade binucléaire après la méiose I.





**Figure 6. Locus de *spo6*.** Le gène *spo6* est situé sur le brin sens du chromosome II de *Schizosaccharomyces pombe*. Le gène qui lui fait face sur le brin anti-sens est le gène *lam2*. Figure adaptée de [www.pombase.org](http://www.pombase.org).

La petite fraction de cellules pouvant cependant passer le stade de la seconde division méiotique montraient une séparation incomplète de leur chromatides sœurs résultant probablement d'un défaut du fuseau mitotique <sup>28</sup>. Une seconde étude <sup>29</sup> a ensuite démontré que la protéine Spo6 interagissait avec la protéine Spo4 pour former un complexe méiotique. Les cibles de ce complexe sont cependant encore inconnues. Le locus de *spo6* est représenté en figure 6 (Fig. 6).

L'étude du rôle de Dus4 dans l'expression du gène *spo6* constitue un des objectifs de ce mémoire.

#### I.4 Méthodes de détection des modifications post-transcriptionnelles

Comme souligné précédemment, il est primordial de pouvoir établir la distribution d'une modification épitranscriptomique d'intérêt au sein du transcriptome. La seconde partie de ce mémoire concerne donc le développement d'une méthode de détection de la dihydrouridine sur l'ensemble des transcriptomes de *S. pombe*, *E. coli* et *C. elegans*. Avant de fixer plus précisément les objectifs de cette deuxième partie, quelques techniques actuelles des études épitranscriptomiques des modifications post-transcriptionnelles les plus étudiées actuellement sont ici présentées. A noter que le développement de ces techniques est toujours à un stade préliminaire et que par conséquent ce domaine de recherche est en rapide expansion.

##### I.4.1 Historique

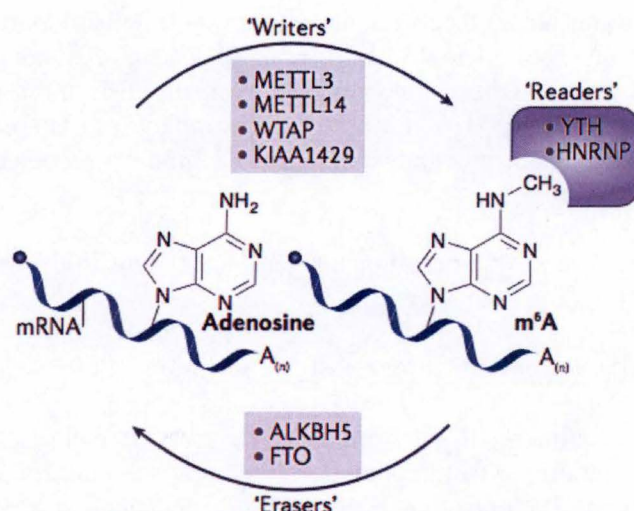
Les nucléotides modifiés ont d'abord été découverts en utilisant la spectrométrie de masse et d'autres techniques de mesures analytiques. L'apparition des techniques de « next-generation sequencing » a ensuite permis de déterminer la présence du nucléotide modifié au sein de séquences de nucléotides, ce qui rend donc possible d'estimer sa distribution dans le transcriptome <sup>7</sup>. Ces cinq dernières années, quelques études épitranscriptomiques pionnières ont émergé, fournissant des données précises sur la position de modifications post-transcriptionnelles au sein du transcriptome <sup>8</sup>.

Le premier essai d'étude large échelle d'une modification post-transcriptionnelle date de 2012 <sup>30</sup> et concerne le m<sup>6</sup>A. Deux équipes indépendantes ont mis au point deux méthodes basées sur une immunoprécipitation spécifique du m<sup>6</sup>A méthylé (m6A-seq <sup>31</sup> et MeRIP-seq <sup>32</sup>), permettant d'établir une liste d'environ 10 000 positions de méthylations de l'adénine. Depuis, le développement de nouvelles approches pour étudier l'impact des modifications sur la structure de l'ARN *in vivo* permet de révéler la plasticité, complexité et fonctionnalité des structures de l'ARN <sup>9</sup>.

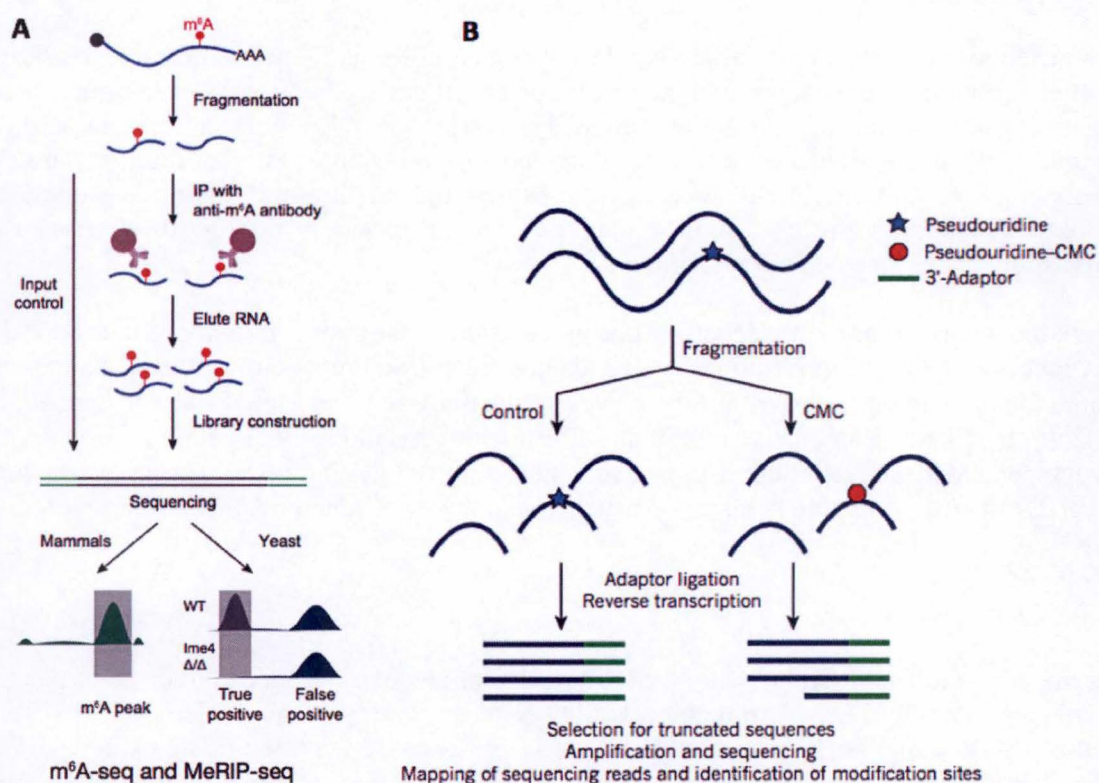
##### I.4.2 Généralités

Les majorités des modifications de l'ARN ne peuvent pas être détectées directement par les techniques de séquençage à haut-débit actuelles. Souvent, la rétro-transcription de l'ARN (qui produit l'ADNc qui lui sera séquençé) efface la présence de ces modifications et les rend indiscernables des bases « classiques » de l'ARN <sup>30</sup>. D'autres stratégies doivent donc être combinées aux techniques de séquençage à haut débit. La plupart des méthodes de détection actuelles des modifications d'ARN se base sur des techniques d'immunoprécipitation de la modification. Dans certains cas, elles sont suivies par la détection d'une empreinte chimique spécifique des nucléotides modifiés qui bloquera la rétro-transcription.





**Figure 7. Les writers, erasers et readers du m<sup>6</sup>A.** Writers actuellement caractérisés : methyltransferase-like3 (METTL3), METTL14, Wilms tumour 1-associated protein (WTAP), KIAA1429 (ces protéines forment un complexe, et c'est METTL3 qui porte l'activité catalytique). Erasers actuellement caractérisés : fat mass and obesity-associated protein (FTO), alkB homologue 5 (ALKBH5). Readers actuellement caractérisés : protéines membres de la famille des protéines contenant un domaine d'homologie YT521 (YTH) et protéines membres de la famille des ribonucléoprotéines hétérogènes nucléaires (HNRNP) <sup>6</sup>.



**Figure 8. Exemples de méthodes de séquençage du transcriptome.** A) Technique de détection du m<sup>6</sup>A (parmi d'autres), basée sur une immunoprécipitation des ARN contenant la modification <sup>30</sup>. B) Technique de détection de la pseudouridine basée sur la modification chimique de la pseudouridine qui mènera à un arrêt de la rétro-transcription. Au séquençage, la détection de la position de la modification se fait grâce au fait que les ADNc obtenus sont tronqués en 3' là où se trouvent la modification <sup>11</sup>.



Enfin, après séquençage du cDNA produit, une approche bio-informatique est nécessaire pour détecter la présence de la modification à l'échelle du nucléotide. La difficulté actuelle en épitranscriptomique est que pour chaque modification, une technique spécifique de la modification est nécessaire. Et pour rappel, 170 modifications post-transcriptionnelles ont pour le moment été répertoriées.

Deux exemples de modifications post-transcriptionnelles et techniques de séquençage associées sont ici présentés, avant de se focaliser sur la méthode de détection de la dihydrouridine. Le premier exemple concerne le m<sup>6</sup>A, car il s'agit de la modification post-transcriptionnelle la mieux caractérisée à l'heure actuelle dans le monde de l'épitranscriptomique. Le second exemple concerne la pseudouridine, dont l'étude a inspiré le développement du D-seq, technique de séquençage de la dihydrouridine, au laboratoire.

#### I.4.3 Exemples

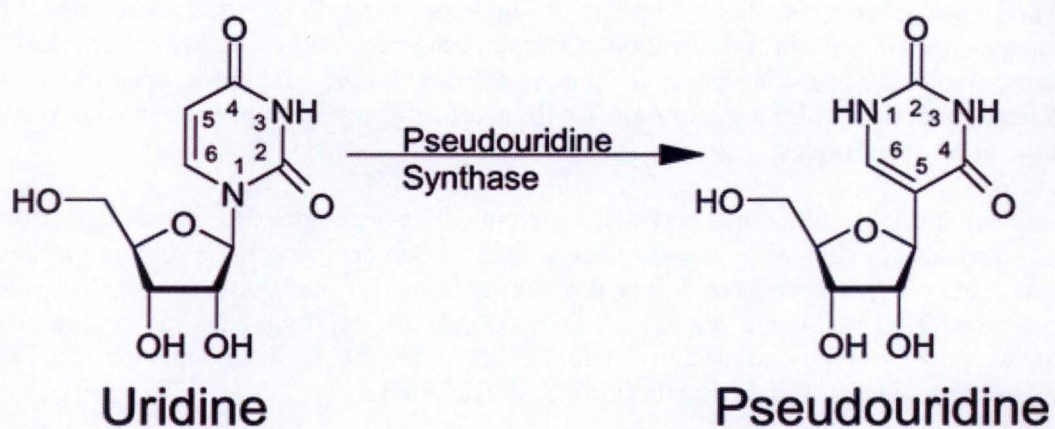
- N<sup>6</sup>-méthyladénosine (m<sup>6</sup>A)

Le m<sup>6</sup>A serait la modification de l'ARNm la plus abondante. Sa première détection date des années 1970<sup>4</sup> et les premières données sur sa distribution à large échelle datent de 2012<sup>31</sup>. Le m<sup>6</sup>A est présent dans plus de 25% des transcrits humains, avec une présence accrue au sein des longs exons, près des codons stop et dans la 5'UTR (5'untranslated region)<sup>6</sup>.

L'effet structurel le plus important provoqué par le m<sup>6</sup>A est un changement de la conformation anti du lien adénine-uridine en une conformation syn. Cela provoque une modification de la structure secondaire de l'ARNm et aboutit à une augmentation de l'accessibilité aux protéines de liaisons à l'ARNm<sup>8</sup>. Les writers du m<sup>6</sup>A sont assemblés sous forme d'un complexe constitué de « methyltransferase like » (methyltransferase like 3 /METTL3, methyltransferase like 14/ METTL14) et du WTAP (Wilms' tumor 1-associated protein). Les deux erasers actuellement référencés du m<sup>6</sup>A sont les protéines FTO (fat mass and obesity-associated protein) et ALKBH5 (alkylation repair homologue protein 5)<sup>2</sup> (**Fig. 7**). Biologiquement, le m<sup>6</sup>A influence presque tous les stades du métabolisme de l'ARNm. Des études récentes ont mis en évidence l'importance du m<sup>6</sup>A dans les machineries de transcription, traduction et d'import-export du noyau<sup>33</sup>. Il a aussi été proposé que le m<sup>6</sup>A favorisait la différenciation cellulaire, cela a été démontré chez la drosophile la souris et les cellules souches humaines embryonnaires<sup>33</sup>. De plus, une perturbation de la machinerie du m<sup>6</sup>A menant à la diminution de la méthylation de l'ARNm serait aussi impliqué dans certains cancers (cancer du sein, carcinome hépatocellulaire, leucémie)<sup>33</sup>. Le m<sup>6</sup>A possède donc un rôle important dans la dynamique du contrôle de l'expression des gènes, fonction qui dépend largement des conditions environnementales cellulaires (stress, choc thermique par exemple).

Dans les deux techniques de détection du m<sup>6</sup>A de 2012 citées précédemment (m<sup>6</sup>A-seq<sup>31</sup> et MeRIP-seq<sup>32</sup>), l'ARN messenger purifié est fragmenté en 100-150 nucléotides et immunoprécipité par un anticorps spécifique du m<sup>6</sup>A. Les fragments d'ARN enrichis en m<sup>6</sup>A passent ensuite par la construction d'une librairie de séquençage et du séquençage à haut débit en tant que tel<sup>30</sup>. Plus tard, une méthode de détection du m<sup>6</sup>A de résolution plus élevée a été appliquée à l'étude de *Saccharomyces cerevisiae* en méiose. Dans cette étude, une souche déficiente en m<sup>6</sup>A-méthyltransférase est utilisée comme contrôle pour éliminer les faux positifs (**Fig. 8**). Récemment, des stratégies de crosslinking aux UV ont été utilisées pour permettre de détecter le m<sup>6</sup>A à l'échelle du nucléotide dans les cellules humaines<sup>30</sup>.





**Figure 9. Conversion de l'uridine en pseudouridine.** Le writer est la pseudouridine synthase (PUS) qui isomérisé l'uridine par clivage de la base et du ribose suivi d'un rattachement pour générer un lien glycosidique C-C<sup>34</sup>.



Bien que les avancées dans la résolution de la détection du m<sup>6</sup>A ne sont pas négligeables, la stœchiométrie des sites du m<sup>6</sup>A reste peu claire. Des spike in contrôle (fragments d'ARN contenant un pourcentage de la modification concernée) sont ajoutés comme standards dans chaque échantillon pour remédier à cette limitation. Le niveau de m<sup>6</sup>A par gène peut ensuite être quantifié grâce à un ratio normalisé de ces contrôles<sup>30</sup>.

Toutes les méthodes actuelles d'étude du m<sup>6</sup>A sont donc basées sur l'utilisation d'anticorps. Cependant, des études récentes ont montré que les anticorps utilisés pouvaient avoir des biais intrinsèques, soit basés sur la séquence de l'ARN, soit sur leurs structures secondaires<sup>30</sup>.

- Pseudouridine ( $\psi$ )

Aussi appelée le cinquième nucléotide, la pseudouridine représente 0,2 à 0,6% de tous les résidus d'uridine de la cellule. Elle jouerait un rôle dans la structure secondaire de l'ARN<sup>33</sup>. De nombreux autres effets biologiques lui ont aussi déjà été attribués. Des expériences *in vitro* ont montré que les ARN pseudouridylés possédaient un temps de vie augmenté<sup>8</sup> et que la suppression de son writer (pseudouridine synthase PUS7) diminuait l'abondance des ARN contenant la pseudouridine (rôle dans la stabilité des transcrits)<sup>33</sup>. Une autre étude a montré que la pseudouridine empêchait l'épissage d'ARN pré-messager en affectant le facteur auxiliaire U2 (U2AF) (rôle dans l'épissage)<sup>2</sup>. La pseudouridine diminuerait aussi la traduction mais des études *in vivo* sont encore nécessaires pour le confirmer (rôle dans la traduction).<sup>2</sup>

Comparativement à l'uridine, la pseudouridine possède un lien hydrogène donneur supplémentaire et un lien C-C plus stable<sup>30</sup> (**Fig. 9**). Cependant, ces caractéristiques ne changent pas les propriétés d'appariements Watson-Crick de la pseudouridine, ce qui rend impossible de distinguer la pseudouridine de l'uridine par séquençage direct.

Pour contourner cette difficulté, la pseudouridine a été modifiée chimiquement en une structure qui provoque l'arrêt de la rétro-transcription. Et ce en utilisant du CMC (N-cyclohexyl-N'- $\beta$ -(4-methylmorpholinium)ethylcarbodiimide) qui modifie la pseudouridine au niveau du N3, l'azote qui sert typiquement de donneur de lien hydrogène dans les appariements Watson-Crick<sup>8</sup>. La pseudouridine modifiée peut donc ainsi bloquer la reverse transcriptase et cela mène à l'obtention de cDNA prématurément tronqués en 3'<sup>30</sup> (**Fig. 8**). Cette méthode a permis d'identifier entre 50 et 100 sites de pseudouridine chez les ARN messagers de la levure et de 100 à 400 sites chez les ARN messagers de l'humain<sup>30</sup>. L'étude de *Schwartz et al*<sup>34</sup> de 2014 qui a mis au point la technique de  $\psi$ -seq s'avère importante pour la réalisation de la deuxième partie de ce mémoire, car c'est cette technique qui a inspiré la méthode mise au point et décrite plus tard dans cette introduction pour détecter la dihydrouridine.

#### I.4.4 Difficultés

Les techniques utilisant l'immunoprécipitation de l'ARN dépendent beaucoup de la qualité de l'anticorps. Un anticorps sensible et spécifique apportera un haut niveau d'informations comparativement à des contrôles et inversement<sup>7</sup>. Actuellement, des anticorps adaptés et spécifiques sont disponibles uniquement pour le m<sup>6</sup>A, le m<sup>1</sup>A, la 5-méthylcytosine et la 5-hydroxyméthylcytosine<sup>35</sup>. La production d'anticorps est typiquement réalisée par l'immunisation d'un animal chez qui l'expression d'anticorps polyclonaux spécifiques est induite<sup>35</sup>. Cependant, le développement d'anticorps *in vivo* ne permet pas de réaliser une contre sélection et nécessite une détection à grande échelle des anticorps spécifiques pour les discerner des anticorps peu spécifiques<sup>35</sup>.



De plus, les méthodes basées sur l'utilisation d'anticorps demandent une grande quantité d'ARN, ce qui exclut l'investigation possible d'échantillons d'ARN qui sont obtenus à faible dose, par exemple à partir de neurones spécifiques ou de biopsies de patients.

Les méthodes non basées sur l'utilisation d'anticorps (par exemple : traitement de la pseudouridine au CMC) ont aussi leurs inconvénients. Que ce soit une dégradation accrue de l'ARN provoquée par le traitement chimique, une résistance au traitement chimique, ou une spécificité relative du traitement chimique<sup>7</sup>.

Un challenge supplémentaire en épitranscriptomique est la quantification de la fréquence des modifications d'ARN. Sans cela, les rôles physiologiques et fonctionnels de la modification d'ARN sont difficiles à déterminer, d'où l'importance des quantifications contrôles lors de l'analyse des échantillons.

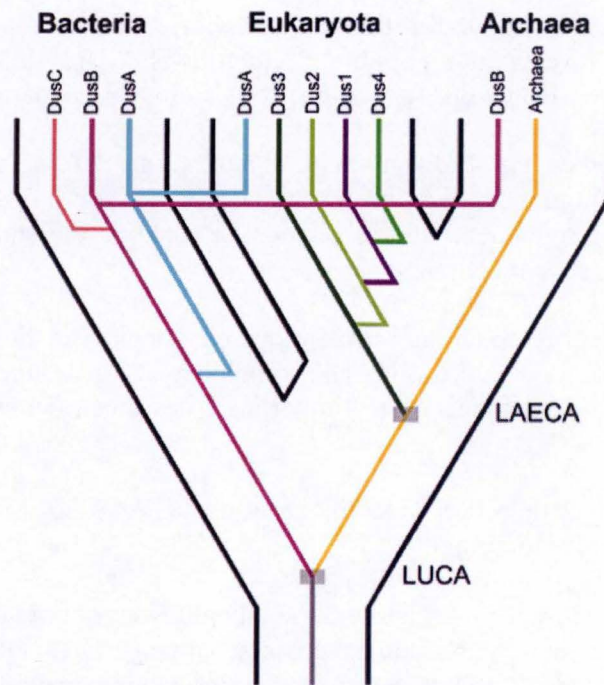
#### I.5 Expérience préalable de détection de la dihydrouridine chez *Schizosaccharomyces pombe* ou « D-seq » et suite du projet

Au laboratoire, dans le cadre d'un projet de thèse en épitranscriptomique (Olivier Finet), une technique de détection de la dihydrouridine a été mise au point. Cette technique se base sur la combinaison d'une méthode d'arrêt de la rétro-transcription via une modification chimique de la dihydrouridine et une technique de séquençage complémentaire de cette méthode pour pouvoir déterminer la distribution de la dihydrouridine à l'échelle du transcriptome de *S. pombe*. C'est ce qui sera appelé du « D-sequencing ou D-Seq ».

Aucun anticorps spécifique de la dihydrouridine n'étant actuellement présent sur le marché, il n'était donc pas possible d'envisager de se baser sur une technique d'immunoprécipitation de l'ARN pour isoler la dihydrouridine. Un ancien protocole de marquage des dihydrouridines par de la rhodamine avait néanmoins été publié<sup>36</sup>. Sur cette base, il a été envisagé que la dihydrouridine chimiquement modifiée puisse engendrer un arrêt de rétro-transcription (technique de « rhodamine labelling », voir plus loin).

L'étude de *Schwartz et al*<sup>34</sup> a été utilisée comme base pour le séquençage des ARN ayant engendré un arrêt de la rétro-transcription après modification chimique de la dihydrouridine. En effet, *Schwartz et al*<sup>34</sup> se basent sur l'induction d'une modification chimique de la pseudouridine pour causer l'arrêt de la rétro-transcription lors de la synthèse d'ADN complémentaire réalisée lors de la préparation de librairie à envoyer au séquençage (voir précédemment). Suite à l'arrêt de la rétro-transcription, des fragments d'ADNc de différentes tailles sont obtenus.

Pour analyser les données obtenues après séquençage des ADNc, deux paramètres sont utilisés. Premièrement le D-ratio (en référence au  $\psi$ -ratio de *Schwartz et al*<sup>34</sup>), qui correspond à la proportion de fragments séquencés se terminant à une position donnée par rapport à l'ensemble des fragments pour lesquels la rétro-transcription ne se sera pas arrêtée à cette position donnée. Ensuite, le D-fold change (D-fc, en référence au  $\psi$ -fc de *Schwartz et al*<sup>34</sup>) qui correspond au  $\log_2$  du D-ratio des échantillons traités (ayant subi la modification chimique de la modification post-transcriptionnelle) divisé par le D-ratio des échantillons non traités. L'ajout de spike-in (molécule d'ARN synthétique composé d'un pourcentage connu et constant en pseudouridine) permet de confirmer la capacité de la technique à déterminer les sites  $\psi$ .



**Figure 10. Arbre phylogénétique potentiel des dihydrouridine synthases.** LAECA est l'ancêtre commun des archées et eucaryotes. LUCA est l'ancêtre commun des bactéries, eucaryotes et archées. Les Dus bactériennes dériveraient de DusB et les Dus eucaryotiques de Dus3<sup>8</sup>.



Dans l'étude de *Schwartz et al*<sup>34</sup>, seulement 43% des reads sur l'ensemble des reads de spike in pseudouridylés étaient lus comme bien pseudouridylés, suggérant une efficacité modérée du traitement chimique, ou de l'arrêt de la rétro-transcription<sup>34</sup>.

Le D-seq d'un premier extrait d'ARN de *S. pombe* et l'analyse des résultats ont été réalisés (voir partie « III) Résultats »), ce qui nous a amenés à transposer la méthode de D-seq à deux autres espèces : *Escherichia coli* et *Caenorhabditis elegans*. Il s'agit du deuxième objectif de ce mémoire : préparer le D-Seq d'*E. coli* et *C. elegans*, dans le but de comparer les résultats avec les premiers résultats obtenus lors du D-seq de *S. pombe* effectué par Olivier Finet.

En effet, la comparaison d'espèces phylogénétiquement peu apparentées pourra fournir des éclaircissements sur de nombreux aspects de la biologie des dihydrouridines<sup>37</sup>. La comparaison de génomes permet d'appréhender la façon dont les gènes évoluent et comment l'évolution à un niveau moléculaire révèle l'adaptation et l'évolution phénotypique de l'organisme. On parle dans ce cas de « génomique comparative »<sup>38</sup>. Dans ce cas-ci, étant donné que c'est l'épitranscriptome qui serait comparé et non le génome, l'on pourrait parler « d'épitranscriptomique comparative ».

Une application possible est de rechercher si une séquence (ici particulièrement : une modification post-transcriptionnelle) retrouvée dans un ou plusieurs génomes est conservée, afin de prédire son possible rôle fonctionnel. Une autre application possible est de rechercher si une séquence (ici : une modification post-transcriptionnelle) prétendument conservée est présente dans toutes les espèces examinées ou s'il existe une perte ou un gain de ce gène (modification post-transcriptionnelle), pouvant alors être associé à de la divergence phénotypique<sup>38</sup>.

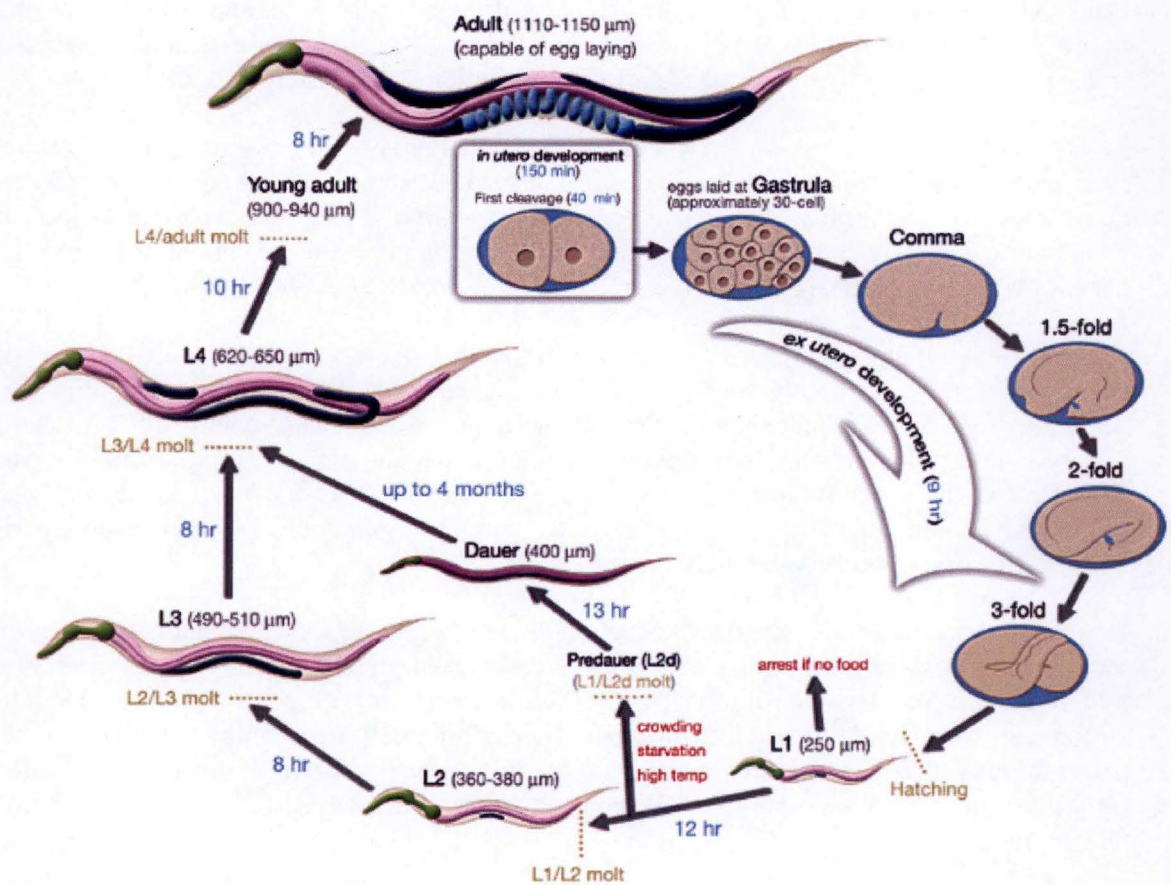
## I.6 Choix de deux autres espèces modèles

En choisissant ces deux espèces modèles, les données du D-seq pourraient alors être comparées entre un organisme procaryote (*Escherichia coli*), un organisme eucaryote unicellulaire (*Schizosaccharomyces pombe*) et un organisme eucaryote pluricellulaire (*Caenorhabditis elegans*). Des études de comparaison phylogénétique ont déjà été réalisées concernant les dihydrouridine synthases<sup>16</sup>. Dus3 serait l'enzyme eucaryote ancestrale, à partir de laquelle les autres enzymes auraient été générées par duplications. Et Dus B serait l'ancêtre des Dus bactériennes<sup>16</sup> (**Fig. 10**). Aucune comparaison phylogénétique concernant la dihydrouridine en tant que telle n'existe à ce jour.

### I.6.1 *Escherichia coli*

*Escherichia coli* est une bactérie Gram-négative naturelle de l'intestin des mammifères découverte en 1885. Elle ne sporule pas et est anaérobie facultative. Elle mesure environ 2 µm de long pour un diamètre de 0,25 à 1 µm. Pour rappel, la taille de *S. pombe* en croissance varie de 8 à 15 µm de long. Elle se multiplie par fission binaire, préférentiellement à une température de 37°C, les cellules filles étant génétiquement identiques aux cellules parentes. Si le nombre de cellules filles reste supérieur au nombre de cellules mères, la croissance d'*E. coli* est dite exponentielle.<sup>58</sup> *E. coli* constitue une grande variété de populations qui possèdent un haut degré de diversité phénotypique et génétique. D'ailleurs, seulement 2% du génome semble être commun parmi toutes les souches d'*E. coli*. Un génome individuel d'*E. coli* contient environ de 4 000 à 5 500 gènes. Cependant, plus de 16 000 séquences de gènes ont déjà été rapportées.





**Figure 11. Cycle de vie de *Caenorhabditis elegans*.** 0 min correspond au moment de la fertilisation. La première division apparaît à 40 min après la fertilisation. La longueur de l'animal à chaque stade est indiquée, tout comme la taille moyenne de chaque stade larvaire (figure provenant du wormatlas).



Cela reflète le fait qu'à peu près deux tiers du pan-génome d'*E. coli* dérive d'autres espèces, et ce par transfert horizontal de gènes <sup>58</sup>.

En sciences, *E. coli* est largement utilisée comme organisme modèle dans les études de microbiologie bactérienne. La bactérie a aussi été utilisée pour comprendre la conjugaison bactérienne et la génétique des phages <sup>58</sup>.

En ingénierie génétique, *E. coli* est très utile grâce à ses outils de conjugaison et de transduction qui permettent aux scientifiques de créer des souches modifiées génétiquement pour ainsi mieux comprendre les processus génétiques et les étudier <sup>58</sup>. Toutes les souches d'*E. coli* utilisées en laboratoire dérivent soit d'*E. coli* B ou *Escherichia coli* K-12.

Concernant la dihydrouridine, quelques études ont été réalisées chez *E. coli*, et plus particulièrement sur les dihydrouridine synthases (pour rappel : DusA, DusB et DusC). Une étude importante pour la réalisation de ce mémoire a été l'étude de Bishop et al de 2002 <sup>39</sup> qui a identifié les membres de la famille des dihydrouridine synthases des ARN de transfert chez *E. coli*. En se basant notamment sur de la génomique comparative, trois gènes correspondant aux enzymes ont été identifiés : *yjbN* pour la protéine DusA, *yhdG* pour la protéine DusB et *yohI* pour la protéine DusC. Des souches knock out pour ces gènes ont été créées et utilisées afin de déterminer les différentes positions de dihydrouridine sur l'ARN de transfert (tRNA<sup>trp</sup> ici). Les souches créées dérivent de la souche *E. coli* K12 <sup>39</sup>.

Pour pouvoir réaliser le D-seq sur une souche wild type et une souche contrôle délétée en gènes codant pour la dihydrouridine synthase et donc normalement ne contenant aucune dihydrouridine, les souches wild type et *ΔyjbN ΔyhdG ΔyohI* de l'étude de Bishop et al <sup>39</sup> ont été utilisées.

#### 1.6.2 *Caenorhabditis elegans*

*Caenorhabditis elegans* est un eucaryote invertébré multicellulaire qui mesure un mm de long à l'âge adulte. C'est dans les années 1960 que l'utilisation de ce vers en tant qu'organisme modèle commence à apparaître, avec notamment l'isolement de la souche N2 (wild type). Il a particulièrement été utilisé pour étudier des mécanismes tel que l'apoptose, les ARN interférents, les voies de signalisation de résistance au stress, les interactions hôtes-pathogènes, et les bases génétiques et moléculaires du développement <sup>40</sup>.

*C. elegans* grandit préférentiellement à des températures comprises entre 15 et 25°C sur milieu gélosé ou en culture liquide, et nourrit de bactéries *E. coli*. *C. elegans* est hermaphrodite autofécondant. Il existe néanmoins des mâles en faible fréquence dans la nature (de 0,1 à 0,2 %). Ceux-ci résultent d'une non-disjonction de la paire de chromosomes sexuels (XX) au cours de l'ovogénèse. Les mâles ont donc un seul chromosome sexuel (X0) et sont distinguables morphologiquement des hermaphrodites : ils sont plus petits et plus fins. La reproduction est possible entre les mâles et les hermaphrodites. La croissance du vers passe par 4 stades larvaires (L1, L2, L3, L4) (**Fig. 11**). Une particularité est qu'à la fin du stade larvaire L2, le vers peut entrer dans un état de croissance « arrêtée » dite « larve dauer » et ce si les conditions environnementales ne sont pas favorables pour le vers (source : « wormatlas »). Chaque adulte pond environ 200 œufs. A 20°C, il faut environ 3 jours pour qu'un œuf pondu par un adulte devienne adulte à son tour. En conditions standards de laboratoire et sans mutations, un vers peut vivre en moyenne 20 jours.



*C. elegans* possède de nombreuses constantes, que ce soit anatomiquement ou biologiquement depuis les parents jusqu'aux descendants, ce qui en fait un organisme modèle d'autant plus étudié<sup>40</sup>.

Concernant le génome du vers, ses 6 chromosomes ont été séquencés pour la première fois en 1998<sup>40</sup>. *C. Elegans* possède 20 000 gènes en moyenne et 74% des protéines humaines possèdent des homologues chez le vers. *C. elegans* contient la majorité des types d'ARN présents chez les eucaryotes, plus de 1300 de ces gènes ont été pour le moment identifiés<sup>41</sup>. Toutes les caractéristiques génétiques connues de *C. elegans* sont reprises sur le site de référence [www.wormbase.org](http://www.wormbase.org).

Concernant la dihydrouridine, aucune étude sur *C. elegans* n'a encore été réalisée à ce jour. *C. elegans* étant un eucaryote pluricellulaire, l'on pourrait cependant s'attendre à ce que l'organisme dispose des 4 dihydrouridine synthases (Dus1, Dus2, Dus3, Dus4), étant donné les résultats obtenus lors des études sur la comparaison phylogénétique des dihydrouridine synthases<sup>16</sup>. Ne disposant pas de souche contrôle, il est prévu ici lors du D-seq d'uniquement séquencer une souche wild type (vers N2).



## II) Objectifs



## II) Objectifs

Ce mémoire s'imbrique donc dans le projet de thèse d'Olivier Finet, concernant l'étude de la dihydrouridine, modification post-transcriptionnelle encore peu caractérisée. Deux choses sont importantes lors de l'étude d'une modification post-transcriptionnelle : comprendre sa distribution au sein du transcriptome, et établir son rôle biologique.

En ce qui concerne le rôle biologique potentiel de la dihydrouridine, la première partie du mémoire découle d'un constat antérieur : celui d'un défaut de sporulation (gamétogenèse) observé dans une souche de *Schizosaccharomyces pombe* délétée pour le gène codant pour l'enzyme dihydrouridine synthase 4 (Dus4). Le but de cette première partie du mémoire est donc d'investiguer les causes de ce phénotype pour comprendre le rôle éventuel de la dihydrouridine dans la sporulation chez *Schizosaccharomyces pombe*.

En ce qui concerne la distribution de la dihydrouridine au sein du transcriptome, la deuxième partie du mémoire s'articule autour de la préparation du « D-seq », une méthode pour établir la distribution des dihydrouridines au sein du transcriptome, dans deux espèces modèles : *Escherichia coli* et *Caenorhabditis elegans*. Au cours de sa thèse, Olivier Finet avait déjà réalisé le D-seq chez la levure de fission *Schizosaccharomyces pombe*. Notre but ici est de transposer les étapes préparatoires à la librairie de séquençage pour deux autres espèces aux caractéristiques biologiques, fonctionnelles et génomiques différentes, dans l'optique d'une comparaison phylogénétique future.



### III) Matériels et Méthodes



### III) Matériels et Méthodes

#### III.1 Implication de la dihydrouridine synthase Dus4 dans la sporulation chez *Schizosaccharomyces pombe*

##### Souches et milieux

Les souches de *Schizosaccharomyces pombe* utilisées dans ce mémoire sont répertoriées dans la table 1.

Toutes les souches sont cultivées en YES (Yeast Extract with Supplements) à 32°C et à une  $DO_{595nm}$  de 0,5 ( $5,5 \times 10^6$  cellules/mL) ou sur YES agar 2% (milieu solide) à 32°C. Afin d'induire la sporulation pour compter les différents types cellulaires, les différentes souches (wild type (WT),  $\Delta dus4$ ,  $\Delta dus4_{BIS}$ ,  $dus4_{DEAD}$ ,  $dus4_{STOP}$ ,  $\Delta spo6$ ,  $dus4::hph^R$ ,  $hst2::ble^R$ ,  $dus4::ura4$ ,  $dus2_{DEAD}$ ,  $\Delta dus4 dus2_{DEAD}$ ) sont striées sur milieu solide MEA (Malt Extract Agar, milieu sur lequel les levures entrent directement en sporulation dû à la privation en azote) deux jours à 25°C (température optimale pour permettre la sporulation). Pour les Northern Blots et qRT-PCRs, les souches wild type,  $\Delta dus4$ ,  $\Delta spo6$ ,  $dus4_{DEAD}$  ont été changées de milieu à partir de MSL (Minimal Sporulating Liquid) vers du MSL-N (Minimal Sporulating Liquid without nitrogen) à 25°C pour induire la sporulation. Pour la sélection des clones sur boîtes, du YES agar 2% (milieu riche) ou de l'EMM (Edinburgh Minimal Medium agar 2%, milieu pauvre) a été utilisé. Le YES a été supplémenté en antibiotiques (généticine, appelée ici kanamycine, à 0.2 µg/mL pour  $kan^R$ ; nourséothricine à 0.1 mg/mL pour  $nat^R$ ; hygromycine B à 0.3 mg/mL pour  $hph^R$ ; phléomycine, appelée ici bléomycine, à 20 mg/mL pour  $ble^R$  ou 5-FOA (5-Fluoroorotic acid) à 1mg/mL).

##### Constructions de souches

Les primers utilisés dans ce mémoire sont répertoriés dans la table 2.

*Note:* travailler avec la levure (organisme unicellulaire) permet de réaliser une amplification PCR directement sur une colonie formée à partir d'une unique cellule. Pour ce faire, les cellules sont chauffées à 100°C pendant 10 minutes et une amplification par ADN polymérase (GoTaq G2 DNA Polymerase, Promega, ref: M784B) est réalisée directement sur l'extrait de cellules suivant les instructions du fabricant. Dans ce mémoire, une telle PCR sera appelée « PCR colonie ».

**$\Delta dus4_{BIS}$  et  $dus4::hph^R$ .** Une PCR Expand (Expand™ High Fidelity PCR System, Sigma-Aldrich 11732650) a été réalisée sur un vecteur pFA6A contenant la cassette  $ble^R$ , avec des primers conçus pour avoir une homologie de séquence avec  $dus4$  (Pombe PCR Primer Programs, [www.bahlerlab.info](http://www.bahlerlab.info)). Le produit PCR purifié a été transformé dans une souche wild type suivant une procédure lithium acétate<sup>42</sup>. Les transformants positifs ont été sélectionnés sur YES + bléomycine ou YES + hygromycine et vérifiés par PCR colonie.

**$dus4_{DEAD}$  et  $dus2_{DEAD}$ .** Pour d'abord construire la souche  $dus4::ura4 ura4-D18$ , une PCR Expand (Expand™ High Fidelity PCR System, Sigma-Aldrich 11732650) a été réalisée sur un vecteur pFA6A contenant la cassette  $ura4$  en utilisant des primers conçus pour avoir une homologie de séquence avec le gène  $dus4$  ou  $dus2$  (Pombe PCR Primer Programs, [www.bahlerlab.info](http://www.bahlerlab.info)). Le produit PCR purifié a été transformé dans une souche portant le marqueur d'auxotrophie  $ura4-D18$  suivant une procédure lithium acétate<sup>1</sup>. Les transformants positifs contenant le gène  $ura4$  au locus  $dus4$  ont été sélectionnés sur EMM (milieu ne contenant pas d'uracile) et YES + 5-FOA 1mg/mL, ce dernier milieu étant toxique pour les clones  $ura^+$ .



Une PCR colonie a été réalisée pour vérifier la présence du gène *ura4* au locus attendu sur les clones positifs. Pour ensuite introduire le gène *dus4* ou *dus2* muté, un fragment de gène (gBlock) contenant deux mutations pour le site catalytique de la protéine Dus4 ou Dus2 (C108A K149A) et une mutation silencieuse correspondant au site de restriction de l'enzyme Sall, a été conçu (Thermo Scientific Fast Digest FD644). Le site Sall a été utilisé pour discriminer le produit PCR *dus4* WT ou *dus2* WT du produit PCR *dus4*<sub>DEAD</sub> ou *dus2*<sub>DEAD</sub>, suite à une simple digestion après PCR, ne nécessitant pas de séquençage. Le gBlock a été amplifié par PCR Phusion (NEB MO5305), purifié et transformé dans une souche *dus4::ura4* ou *dus2::ura4*. Les clones capables de pousser sur YES + 5-FOA ont été sélectionnés. Une seconde PCR colonie a ensuite été réalisée pour vérifier la présence du gène *dus4*<sub>DEAD</sub> dans les clones sélectionnés, suivie d'une restriction avec l'enzyme Sall. Pour perdre le marqueur d'auxotrophie *ura4-D18* de la souche construite, un croisement sur MEA à 25°C pendant 3 jours a été réalisé entre les souches *dus2* ou *dus4::ura4 ura4-D18* et une souche WT. Des dilutions 100x and 1000x ont été mises sur EMM. Les clones qui poussaient sur EMM ont été sélectionnés et une PCR colonie avec les primers du gBlock suivie d'une phase de restriction a été réalisée comme décrit ci-dessus. Le produit a été purifié et vérifié par séquençage.

***Δdus1-2-3- dus4*<sub>DEAD</sub>**. Un croisement a été réalisé entre les souches *dus4*<sub>DEAD</sub> et *Δdus1-2-3-4* (*dus1::nat<sup>R</sup> dus2::kan<sup>R</sup> dus3::hph<sup>R</sup> dus4::ble<sup>R</sup>*) sur MEA à 25°C pendant 3 jours. Les spores furent mises sur YES et sélectionnées sur YES + nourséothricine, kanamycine et hygromycine (sélections des délétions *dus1*, *dus2* and *dus3*) et sur YES + bléomycine (les clones poussant sur ces boîtes portaient la délétion *dus4* et n'étaient donc pas sélectionnés). Des PCR colonies ont été réalisées pour vérifier les loci.

***Δdus4 dus2*<sub>DEAD</sub>**. Un croisement a été réalisé entre les souches *dus2*<sub>DEAD</sub> et *Δdus2 Δdus4* (*dus2::kan<sup>R</sup> dus4::ble<sup>R</sup>*) sur MEA à 25°C pendant 3 jours. Les spores ont été mises sur YES et sélectionnées sur YES + kanamycine et sur YES + bléomycine (les clones poussant sur les boîtes YES + kanamycine portaient la délétion *Δdus2* et n'ont donc pas été sélectionnés). Des PCR colonies ont été réalisées pour vérifier les loci.

***dus4*<sub>STOP</sub>**. Pour d'abord construire la souche *dus4::ura4 ura4-D18*, une PCR Expand (Expand™ High Fidelity PCR System, Sigma-Aldrich 11732650) a été réalisée sur un vecteur pFA6A contenant la cassette *ura4* en utilisant des primers conçus pour avoir une homologie de séquence avec le gène *dus4* (Pombe PCR Primer Programs , [www.bahlerlab.info](http://www.bahlerlab.info)). Le produit PCR purifié a été transformé dans une souche portant le marqueur d'auxotrophie *ura4-D18* suivant une procédure lithium acétate<sup>1</sup>. Les transformants positifs contenant le gène *ura4* ont été sélectionnés sur EMM et YES + 5-FOA 1mg/mL. Un fragment de gène (gBlock) contenant 8 mutations (A28T G37T A46T C92G T158A G499T G584T T707A) correspondant à l'introduction de huit codons STOP prématurés et une mutation silencieuse correspondant au site de restriction de l'enzyme Sall a été conçu (Thermo Scientific Fast Digest ref: FD644). Le gBlock a été amplifié par PCR Phusion (NEB ref: MO5305), purifié et transformé dans une souche *dus4::ura4*. Les clones capables de pousser sur YES + 5-FOA ont été sélectionnés (car ne portant plus de gène *ura*). Une seconde PCR colonie a ensuite été réalisée pour vérifier la présence du gène *dus4*<sub>STOP</sub> dans les clones sélectionnés, suivie d'une restriction avec l'enzyme Sall. Pour supprimer le marqueur d'auxotrophie *ura4-D18* de la nouvelle souche, un croisement sur MEA à 25°C pendant 3 jours a été réalisé entre les souches *dus4::ura4 ura4-D18* et une souche WT. Des dilutions 100x and 1000x furent mises sur EMM. Les clones qui pouvaient pousser sur EMM ont été sélectionnés et une PCR colonie avec les primers du gBlock suivie d'une phase de restriction a été réalisée comme décrit ci-dessus. Le produit a été purifié et vérifié par séquençage.



***dus4*<sub>STOP</sub>-TAP, *dus4*<sub>DEAD</sub>-TAP and *dus4*-TAP.** Les souches ont été tagées par l'ajout d'un TAP lié à une cassette de résistance *nat*<sup>R</sup> au niveau du locus. Une PCR Expand (Expand™ High Fidelity PCR System, Sigma-Aldrich 11732650) a été réalisée sur un vecteur pFA6A contenant le locus *TAP-nat*<sup>R</sup> et en utilisant des primers conçus pour avoir une homologie de séquence avec *dus4*<sub>STOP</sub>, *dus4*<sub>DEAD</sub> ou *dus4* (Pombe PCR Primer Programs, [www.bahlerlab.info](http://www.bahlerlab.info)). Le produit PCR purifié a été transformé dans une souche WT selon une procédure lithium acétate<sup>1</sup>. Les transformants positifs ont été sélectionnés sur YES + nourséothricine et vérifiés par PCR colonie.

### Extraction ARN

Les souches qui ont poussé dans du YES à 32°C jusqu'à une DO<sub>595nm</sub> de 0,5 (5,5 x 10<sup>6</sup> cellules/mL) ont été culotées dans de l'eau DEPC (diéthyl pyrocarbonate) réfrigérée, resuspendues dans 750 µL de TES buffer (10 mM Tris pH7,5; 10 mM EDTA pH8; 0,5% SDS) et immédiatement supplémentées avec 750 µL d'acidic phenol-chloroform (refrigéré, Sigma P-1944). Ensuite, les échantillons ont été incubés pendant une heure à 65°C. Les échantillons ont été placés sur glace et centrifugés pendant 15 min à 14000 rpm à 4°C. 700 µL de la phase aqueuse ont été pris et 700 µL d'acidic phenol-chloroform-IAA (refrigéré, Sigma 77619) y ont été ajoutés. Les échantillons ont été centrifugés 5 minutes à 14000 rpm à 4°C. 650 µL de la phase aqueuse ont été pris et 650 µL de chloroform:IAA (température ambiante, Sigma 25666) y ont été ajoutés. Les échantillons ont été centrifugés 5 min à 14000 rpm à 4°C. 500 µL de la phase aqueuse ont été mélangés avec 1,5 ml de 100% EtOH (-20°C) et 50 µL de 3M NaAc pH5.2. Les échantillons ont été précipités à -20°C overnight (ou à -80°C pendant 30 min). Les échantillons ont ensuite été centrifugés pendant 10 min à 14,000 rpm à 4°C. 500 µL de 70% EtOH (4°C, fait avec de l'eau DEPC) ont été ajoutés au culot obtenu. Les échantillons ont ensuite été centrifugés pendant 1 min à 14,000 rpm à 4°C. La plupart du surnageant a été aspiré et le culot séché à l'air ambiant. 200 µL d'eau DEPC ont été ajoutés et les échantillons ont été incubés pendant 10 minutes à 65°C.

### q-RT PCR

Pour la quantification du gène *spo6*, les différentes souches ont été mises dans du MSL-N (Minimal Sporulating Liquid without nitrogen) après avoir poussé un jour dans du MSL. Pour la quantification du gène *dus4*, les souches ont classiquement été mises dans du YES liquide. L'ARN a été extrait par une extraction au phénol 6, 8 et 10 heures après le changement de milieu. Les ARNs ont été purifiés avec un kit RNeasy (Qiagen, ref: 74104), dilués 10x et quantifiés. Les concentrations ont ensuite été ajustées à 0,05 µg/µL. La rétro-transcription a été réalisée par une succession de trois cycles : 10 minutes à 25°C, 2heures à 37°C et 5 minutes à 85°C. 5 µL du produit cDNA après rétro-transcription ont été dilués 100x et mis dans un mix PCR contenant 12,5 µL de SYBR Green (Biorad 1725124). L'abondance du transcrit *spo6* ou *dus4* a été normalisée sur l'abondance du *housekeeping gene act1*.

### Northern Blot

L'ARN total extrait au phénol (5-25 µg) a été chargé sur un gel d'agarose 1x avec 3,7 % de formaldéhyde et 2µL de bromure d'éthidium. Après électrophorèse, l'ARN a été transféré sur une membrane de nylon et crosslinké 2h à 80°C. Les sondes radioactives ont été préparées avec le Random Primers DNA Labeling System (ThermoFisher 18187-013, brin non spécifique), suivant les instructions du fabricant.



### Marquage de la dihydrouridine par la rhodamine (« rhodamine labelling »)

La dihydrouridine est marquée avec de la rhodamine 110 tel décrit par *Betteridge et al*<sup>43</sup> mais avec les modifications suivantes : le marquage a été fait avec 30 µg d'ARN total et éventuellement resuspendu dans 6 µL d'eau DEPC. Les échantillons R+ font référence à un marquage effectif de la dihydrouridine tandis que les échantillons R- se réfèrent à une condition contrôle dans laquelle la préparation de l'échantillon ne résulte pas en l'incorporation de la rhodamine.

### Primer extension

5 à 25 µg d'ARN total "rhodaminisé" ont été utilisés pour la rétro-transcription lors du primer extension. La réaction a été réalisée avec le kit Invitrogen 18080-093 suivant les instructions du fabricant, en utilisant un primer spécifique de l'ARNt préalablement rendu radioactif avec une T4 polynucléotide kinase (Promega M4101) et du [ $\gamma$ -<sup>32</sup>P] ATP (Perkin NEGoo2A25oUC). Le ladder ADN consiste en 10 primers (de 15 à 60 nucléotides de long) mis à une concentration de 5 µM (0,5 µM chacun) et marqués radioactivement comme décrit ci-dessus. Les échantillons et le ladder ont été chargés sur un gel polyacrylamide 15%.

### Western Blot

Les souches ont poussé à DO<sub>595nm</sub> 0,5 (5,5 x 10<sup>6</sup> cellules/mL) et les protéines ont été extraites suivant un protocole d'extraction NaOH et chargées sur un gel mini protean 4-15% TGX (Biorad, ref: 4561083). Les extraits migrés ont été transférés sur une membrane (Biorad Transfer pack, ref: 1704158) puis incubés avec un anticorps primaire anti-TAP (PAP ; Peroxydase Anti-Peroxydase) pour les souches wild type, *dus4-TAP*, *dus4*<sub>DEAD</sub>-TAP, *dus4*<sub>STOP</sub>-TAP, *spo6-TAP*. Pour le contrôle de charge, la tubuline a été révélée avec un anticorps secondaire anti-mouse, après dé-stripping de la membrane porteuse de l'anticorps anti-TAP. Après lavage, les protéines ont été révélées avec un réactif ECL (électrochimiluminescence). La membrane a finalement été exposée à l'ImageQuant.

### III.2 Préparation d'une analyse épitranscriptomique comparative de la dihydrouridine chez *Schizosaccharomyces pombe*, *Escherichia coli* et *Caenorhabditis elegans*

#### Souches et milieux

##### *Escherichia coli*

Les souches WT (JC8679<sup>12</sup>) et  $\Delta$ *dusABC* (PS5107<sup>12</sup>) ont été cultivées en LB (Lysogeny Broth) à 37°C et à une DO<sub>595nm</sub> située entre 0,6 et 0,8 ou sur LB agar 2% (milieu solide) à 37°C. Pour la vérification des résistances aux antibiotiques (chloramphénicol, kanamycine, spectinomycine), du LB agar 2% supplémenté en antibiotiques (chloramphénicol à 0,03 µg/mL ; kanamycine à 1 µg/mL ; spectinomycine à 0.1 µg/mL) ont été utilisés.

##### *Caenorhabditis elegans*

En culture solide, les vers WT (N2) ont poussés sur du NGM agar (Nematode Growth Media) (NaCl :1,5 g pour 500 mL, Bacto Peptone : 1,25 g pour 500 mL, agar : 8,5g pour 500 mL) + suppléments (pour 500 mL : 500 µL de cholestérol 5 µg/mL, 500 µL de CaCl<sub>2</sub>.2H<sub>2</sub>O 1M, 500 µL de MgSO<sub>4</sub> 1 M, 12,5 mL de KH<sub>2</sub>PO<sub>4</sub> 1M) avec 300 µL de bactéries OP50.



En culture liquide, les vers ont soit poussé dans du S-basal 1x (pour 1 L : 5,8 g de NaCl 0,1M, 50 mL  $\text{KH}_2\text{PO}_4$  pH6 1M, 1mL de cholestérol 5ug/mL, 3 mL de  $\text{MgSO}_4 \cdot 7\text{H}_2\text{O}$  3mM, 4mL de  $\text{CaCl}_2 \cdot 2\text{H}_2\text{O}$  4mM, 10 mL de traces de métaux, 10 mL de K citrate. $\text{H}_2\text{O}$ ) ou du M9 buffer 1x (pour 1 L : 3 g de  $\text{Na}_2\text{HPO}_4$  42,2mM ; 1,5g de  $\text{KH}_2\text{PO}_4$  22 mM, 2,5g de NaCl 85,7 mM, 0,5 mL de  $\text{MgSO}_4$  1mM) en fonction du moment du cycle, et nourris avec des bactéries HB101 (4mL pour 200 mL de milieu de culture).

En partant de 10 vers sur NGM agar + OP50 et après passage en culture liquide après un cycle complet, environ 200 000 vers ont été récoltés au stade L4.

## **Extraction ARN**

### ***Escherischia coli***

Les souches WT (JC8679) et  $\Delta dusABC$  (PS5107) ont été cultivées en LB (Lysogeny Broth) à 37°C et à une  $\text{DO}_{595\text{nm}}$  située entre 0,6 et 0,8. Le protocole d'extraction est le même que pour *Schizosaccharomyces pombe* (voir ci-dessus) avec la différence suivante : les échantillons finaux ont été incubés pendant 15 minutes à 65°C.

### ***Caenorhabditis elegans***

Le protocole d'extraction est le même que pour *Schizosaccharomyces pombe* (voir ci-dessus), après récolte d'environ 200 000 L4 par condition.

## **Bioanalyzer**

1  $\mu\text{L}$  d'ARN à concentration 3,5 ng/ $\mu\text{L}$  et 0,35ng/ $\mu\text{L}$  pour chaque échantillon a été chargé sur la puce du kit Bioanalyzer Agilent RNA 6000 Pico kit. Les étapes et l'analyse ont été réalisées selon les informations du fabricant fournies dans le quick start guide (réf : 5067-1513).

## **Marquage de la dihydrouridine par la rhodamine (« rhodamine labelling »)**

Voir ci-dessus

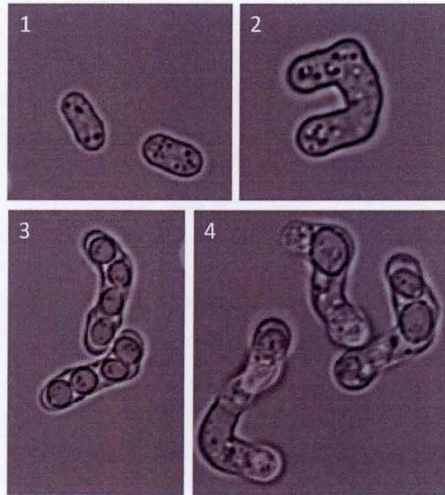
## **Dot blot assay**

Des gouttes des échantillons d'ARN extraits rhodaminisés de 3 $\mu\text{L}$  à concentration 0,5  $\mu\text{g}/\mu\text{L}$  ont été placées sur une membrane de nylon et la fluorescence visualisée (520nm) et quantifiée avec le logiciel Imagquant. Des gouttes contrôle de rhodamine, méthanol et eau ont été utilisés.



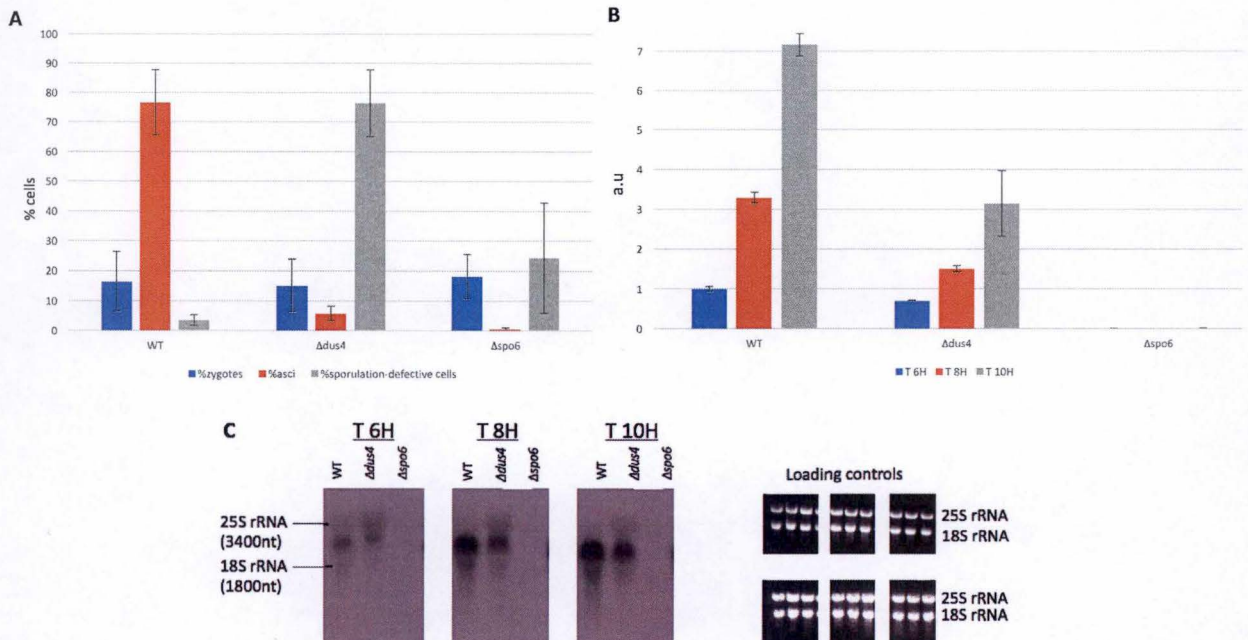
## IV) Résultats





**Figure 12. Observations DIC (microscopie à contraste interférentiel) de *Schizosaccharomyces pombe*.**

1. Deux cellules au stade végétatif 2. Une cellule au stade zygote après conjugaison de deux cellules végétatives. Noter la différence de taille avec les cellules au stade végétatif et la forme caractéristique « en pont » du zygote. 3. Deux asques chacun contenant quatre spores haploïdes 4. Trois asques présentant un défaut de sporulation (« sporulation defective cells » : spores déformées, atrophiées, ou manquantes). Précision : 10  $\mu$ m



**Figure 13. La souche  $\Delta dus4$  présente un défaut de sporulation pouvant être corrélé à un défaut d'expression du gène *spo6*.** A) Pourcentage de zygotes (bleu), asques (orange), asques présentant un défaut de sporulation (gris) 48h après l'induction de la sporulation dans les souches wild type (WT),  $\Delta dus4$  et  $\Delta spo6$ . Les barres d'erreur représentent l'écart-type pour trois répliques biologiques indépendantes (représentant 600 cellules). B) qRT-PCR du gène *spo6* 6h, 8h et 10h après l'induction de la sporulation dans les souches wild type (WT),  $\Delta dus4$  et  $\Delta spo6$ . Les barres d'erreur représentent l'écart-type pour deux répliques techniques, l'ARNm du *house keeping gene act1* a été utilisé pour la normalisation. a.u = arbitrary units (WT 6h = 1) C) Northern Blot du gène *spo6* 6h, 8h et 10h après l'induction de la sporulation dans les souches wild type (WT),  $\Delta dus4$  et  $\Delta spo6$ . Les ARN ribosomiques ont été utilisés comme contrôles de charge, haut : gel d'agarose, bas : membrane de nylon



## IV) Résultats

### IV.1 Implication de la dihydrouridine synthase Dus4 dans la sporulation de *Schizosaccharomyces pombe*

#### IV.1.1 Observation d'un défaut de sporulation dans la souche $\Delta dus4$ corrélé avec une diminution de l'expression du gène *spo6*

Un des premiers objectifs de ce mémoire a été de confirmer l'observation du défaut de sporulation (gamétogenèse) préalablement constaté au laboratoire dans la souche  $\Delta dus4$ . Le défaut de sporulation correspond à la présence de gamètes déformés, atrophiques, ou manquants. Les cellules présentant ce phénotype sont appelées « sporulation defective cells ». Au sein de cette souche, le gène codant pour l'enzyme dihydrouridine synthase 4 (Dus4) a été remplacé par le gène responsable de la résistance à l'antibiotique bléomycine (génotype de la souche : *dus4::ble<sup>R</sup>*). Chez *S. pombe*, le remplacement d'un gène par un marqueur de résistance à un antibiotique permet de sélectionner les clones ayant intégrés le marqueur en lieu et place du gène ciblé, ici : *dus4*.

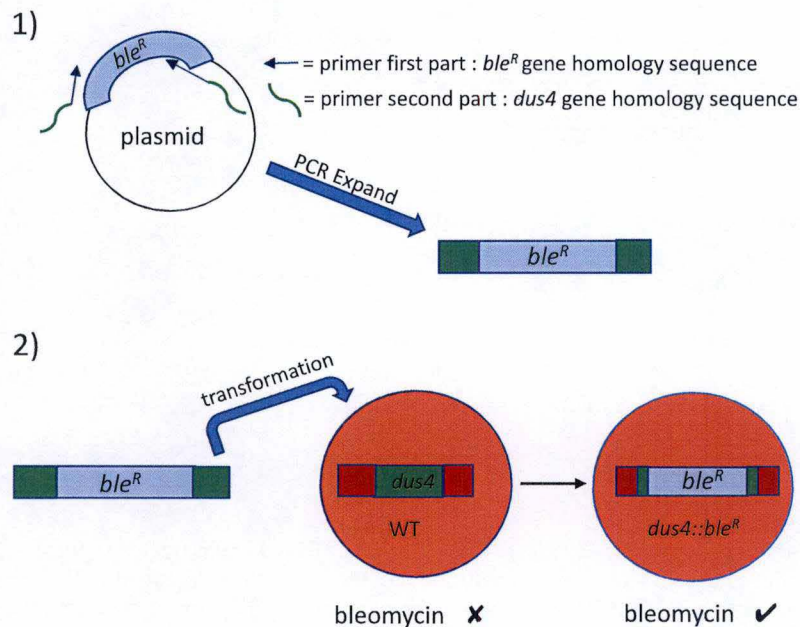
Un milieu carencé en azote permet d'initier la gamétogenèse chez la levure *S. pombe*. Une souche homothallique de génotype *dus4::ble<sup>R</sup>* a été observée après être restée 48 heures à 25°C, ces conditions étant des conditions optimales pour permettre la sporulation. Quatre catégories phénotypiques ont été répertoriées : végétatif – zygote – asque – asque avec un défaut de sporulation (« sporulation defective cell ») (**Fig. 12**). Dans la souche  $\Delta dus4$ , le pourcentage de sporulation defective cells prédomine (76,4%) en comparaison avec une souche wild type (3,5%), confirmant le constat déjà réalisé précédemment au laboratoire (**Fig. 13A**).

Afin de comprendre le mécanisme moléculaire menant à ce phénotype, des profils transcriptomiques avaient été préalablement générés par RNAseq (données non montrées, travail d'Olivier Finet, thèse en cours) pour les souches WT et  $\Delta dus4$ . Ces analyses avaient montré la présence d'une diminution de l'expression du gène de sporulation *spo6* dans la souche  $\Delta dus4$ . Une analyse phénotypique a également été réalisée dans une souche  $\Delta spo6$  (*spo6::nat<sup>R</sup>*) afin de confirmer l'implication de ce gène dans la sporulation. Sur les 24,9 % de cellules atteignant le stade d'asque, 24,5 % étaient des sporulation defective cells (**Fig. 13A**).

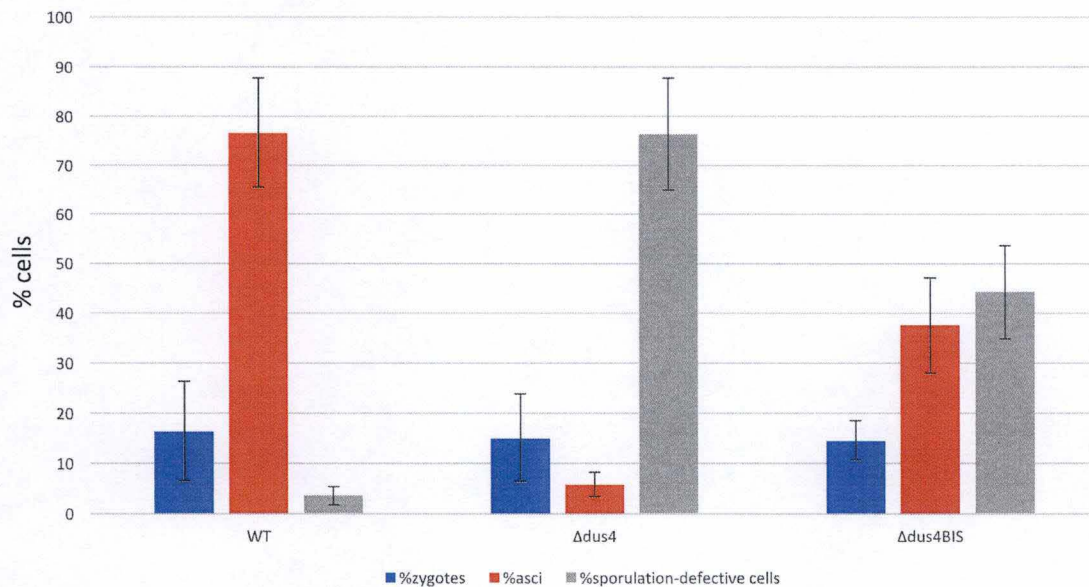
L'impact de la délétion du gène *dus4* sur l'expression du gène *spo6* a donc ensuite été caractérisé par une analyse quantitative et qualitative (qRT-PCR et Northern Blot) de l'ARN messager du gène *spo6*. Un milieu liquide carencé en azote a été utilisé pour l'induction de la sporulation chez les souches wild type,  $\Delta dus4$  et  $\Delta spo6$ . Les ARN totaux ont été extraits 6, 8 et 24 heures après incubation en milieu carencé. L'ARNm du gène *spo6* est non détectable dans la souche  $\Delta spo6$  aussi bien par qRT-PCR (**Fig. 13B**) que par Northern Blot (**Fig. 13C**). En ce qui concerne la souche  $\Delta dus4$ , l'induction de l'expression de *spo6* est fortement diminuée en comparaison avec la souche wild type (**Fig. 13B-13C**), confirmant les données transcriptomiques du RNAseq.

Afin de confirmer de manière plus appuyée l'observation du défaut de sporulation dans la souche  $\Delta dus4$ , une tentative de restauration du phénotype wild type par complémentation avait préalablement été réalisée. Elle consistait en la réintroduction du gène *dus4* dans la souche  $\Delta dus4$  à partir d'un plasmide. De manière très surprenante, la complémentation n'a jamais pu être réalisée avec succès, questionnant l'authenticité de la délétion de *dus4*.



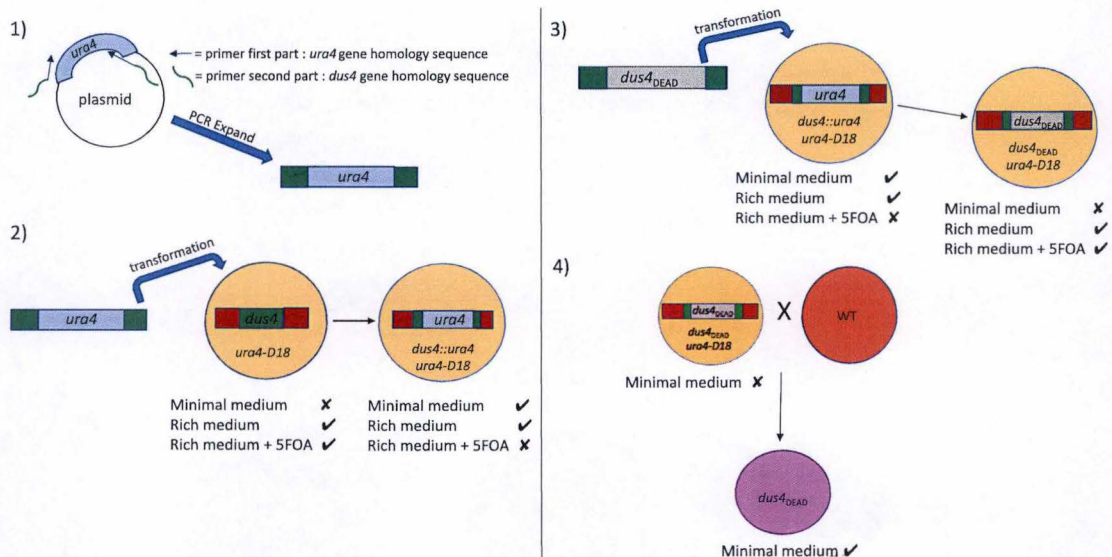


**Figure 14. Construction de la souche  $\Delta dus4_{BIS}$**  1) PCR Expand (PCR à haute-fidélité répliquative) d'un plasmide porteur du gène de la cassette de résistance *ble<sup>R</sup>* avec des primers conçus en laboratoire et qui contiennent une homologie de séquence avec le gène *ble<sup>R</sup>* et le gène *dus4*. 2) Transformation du produit PCR dans une souche wild type contenant le gène *dus4* bordé par ses gènes voisins (rouge). Les clones *dus4::ble<sup>R</sup>* obtenus ont été sélectionnés sur un milieu de sélection (bléomycine) et vérifiés par PCR.



**Figure 15. La souche  $\Delta dus4_{BIS}$  présente un phénotype intermédiaire entre la souche wild type et la souche  $\Delta dus4$ .** Pourcentage de zygotes (bleu), asques (orange), présentant un défaut de sporulation (gris) 48h après l'induction de la sporulation dans les souches wild type (WT),  $\Delta dus4$  et  $\Delta spo6$ . Les barres d'erreur représentent l'écart-type pour trois répliques biologiques indépendantes (représentant 600 cellules).





**Figure 16. Construction de la souche *dus4<sub>DEAD</sub>*.** 1) PCR Expand (PCR à haute-fidélité répliquative) d'un plasmide porteur du gène *ura4* avec des primers conçus contenant une séquence d'homologie avec le gène *ura4* et le gène *dus4*. 2) Transformation du produit PCR dans une souche porteuse du marqueur d'auxotrophie *ura4-D18* (donc incapable de pousser sur milieu défini dépourvu d'uracile). La souche obtenue *dus4::ura4 ura4-D18* a été sélectionnée par sa capacité à pousser sur milieu défini mais pas sur milieu riche + 5-FOA (métabolite toxique sur les souches *ura<sup>+</sup>*). 3) Un fragment de gène (*dus4<sub>DEAD</sub>* en gris) correspondant à la portion du gène *dus4* dans laquelle deux mutations pour l'activité catalytique de la protéine Dus4 ont été conçues (C108A K149A) a été transformé dans la souche *dus4::ura4 ura4-D18* montrée à l'étape 2. La souche obtenue *dus4<sub>DEAD</sub> ura4-D18* était capable de pousser sur milieu riche + 5-FOA mais pas sur milieu défini. 4) Un croisement entre la souche *dus4<sub>DEAD</sub> ura4-D18* et une souche wild type a été réalisé dans le but d'éliminer le marqueur d'auxotrophie *ura4-D18*. Tous les clones obtenus ont été vérifiés par PCR aux différentes étapes de la construction et la souche finale *dus4<sub>DEAD</sub>* a été vérifiée par séquençage.



Nous avons donc réalisé une seconde souche  $\Delta dus4$ , appelée  $\Delta dus4_{BIS}$ . Le but de cette construction était de vérifier la présence du défaut de sporulation en réitérant la délétion du gène *dus4*. Cela aussi dans le but de s'assurer que le défaut observé n'était pas possiblement dû à une altération présente quelque part ailleurs dans le génome de la première souche construite  $\Delta dus4$  (voir **Fig. 14** pour l'explication de la construction  $\Delta dus4_{BIS}$ ).

Une analyse phénotypique a été effectuée sur la souche  $\Delta dus4_{BIS}$  créée après induction de la sporulation. Bien que le phénotype de déficience de sporulation s'avère moins prévalent que dans la souche  $\Delta dus4$  initiale (76,4 %), il reste néanmoins prédominant (44%), ce qui supporte qu'un défaut de sporulation résulte bien du remplacement du gène *dus4* par le marqueur *ble<sup>R</sup>* (**Fig. 15**).

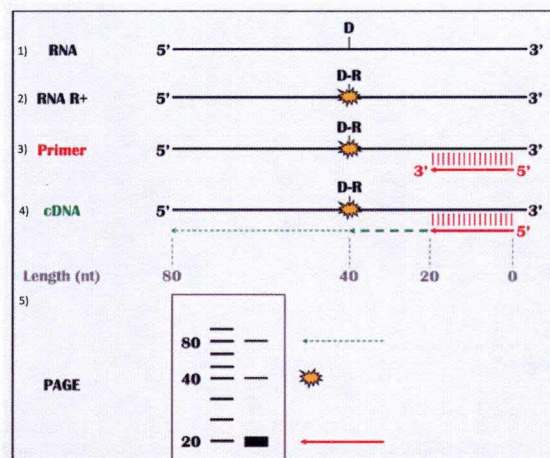
#### IV.1.2 Etude de l'implication de l'activité catalytique de la protéine Dus4 dans le défaut de sporulation observé et corrélé avec une diminution de l'expression du gène *spo6* dans la souche $\Delta dus4$

L'étape suivante dans l'étude du défaut de sporulation observé dans la souche  $\Delta dus4$  a été de comprendre si cela pouvait être dû à l'absence de l'activité catalytique de l'enzyme dihydrouridine synthase 4.

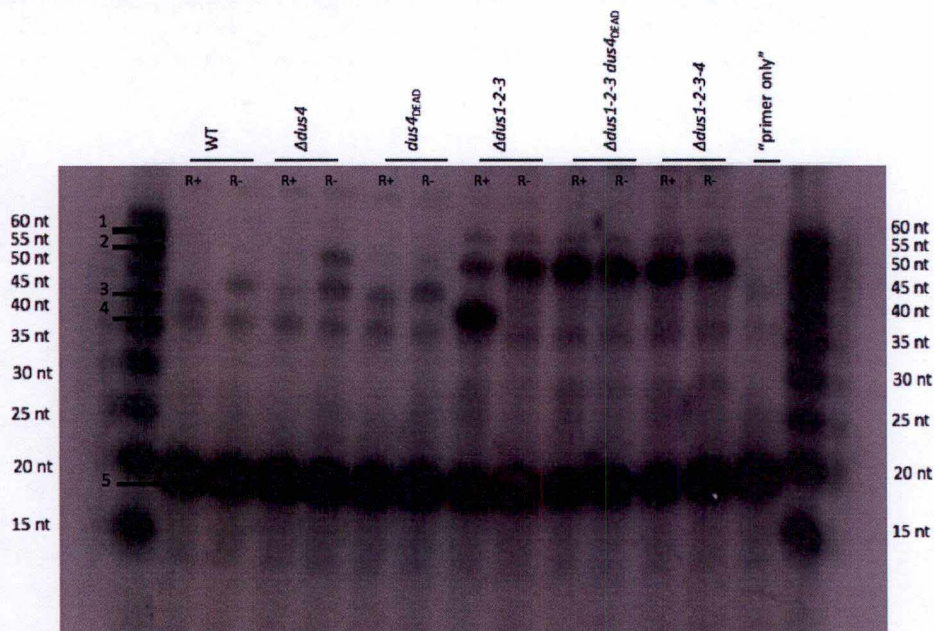
Pour cela, une souche pour laquelle la région du gène *dus4* codant pour l'activité catalytique de la protéine était mutée a été créée. Cette souche appelée *dus4<sub>DEAD</sub>* (pour « dead enzymatic activity ») fut obtenue par une stratégie de sélection en deux étapes (**Fig. 16**). Premièrement, une souche *ura4-D18* auxotrophe à l'uracile a été utilisée. Cette souche est incapable de synthétiser de l'uracile, et donc de pousser sur un milieu sans uracile, suite à la présence d'une forme tronquée du gène *ura4* (gène codant pour l'orotidine-5'-phosphate décarboxylase, enzyme impliquée dans la voie de synthèse de l'uracile<sup>44</sup>). L'auxotrophie a été complémentée en remplaçant le gène *dus4* par le gène *ura4*. La souche obtenue possédait alors le génotype *dus4::ura4 ura4-D18*, et avait récupéré la capacité de pousser sur un milieu sans uracile.

Deuxièmement, la souche construite *dus4::ura4 ura4-D18* a été transformée avec une version synthétique du gène *dus4* (*dus4<sub>DEAD</sub>*) dans laquelle deux codons codant pour des acides aminés essentiels à l'activité catalytique de la protéine avaient été mutés. Les codons à muter (C108A et K149A) ont été choisis grâce à deux études précédentes qui avaient identifié les résidus essentiels pour l'activité catalytique des protéines Dus. Ici, la cystéine en position 108 et la lysine en position 149 furent toutes les deux remplacées par une alanine lors de la création synthétique du gène *dus4<sub>DEAD</sub>*<sup>20, 19</sup>. Par recombinaison homologue, le gène *ura4* se voyait alors remplacé par le gène *dus4<sub>DEAD</sub>*, résultant en une restauration de l'auxotrophie pour l'uracile. Ce phénotype peut difficilement être sélectionné mais une sélection positive peut être appliquée en présence d'acide 5-florootic (5-FOA)<sup>44</sup>, rendu toxique par l'enzyme Ura4 fonctionnelle. Afin de se débarrasser du marqueur d'auxotrophie *ura4-D18*, un croisement a ensuite été réalisé avec une souche wild type et les clones poussant sur milieu sans uracile ont été sélectionnés. Dans ce cas-ci, on a donc sélectionné les clones ayant perdu le chromosome porteur de la délétion *ura4-D18* au profit du gain du chromosome WT. Finalement, la présence des deux codons mutés a été vérifiée par PCR, grâce à un site de restriction présent dans la version synthétique du gène *dus4* utilisée, ainsi que par séquençage.



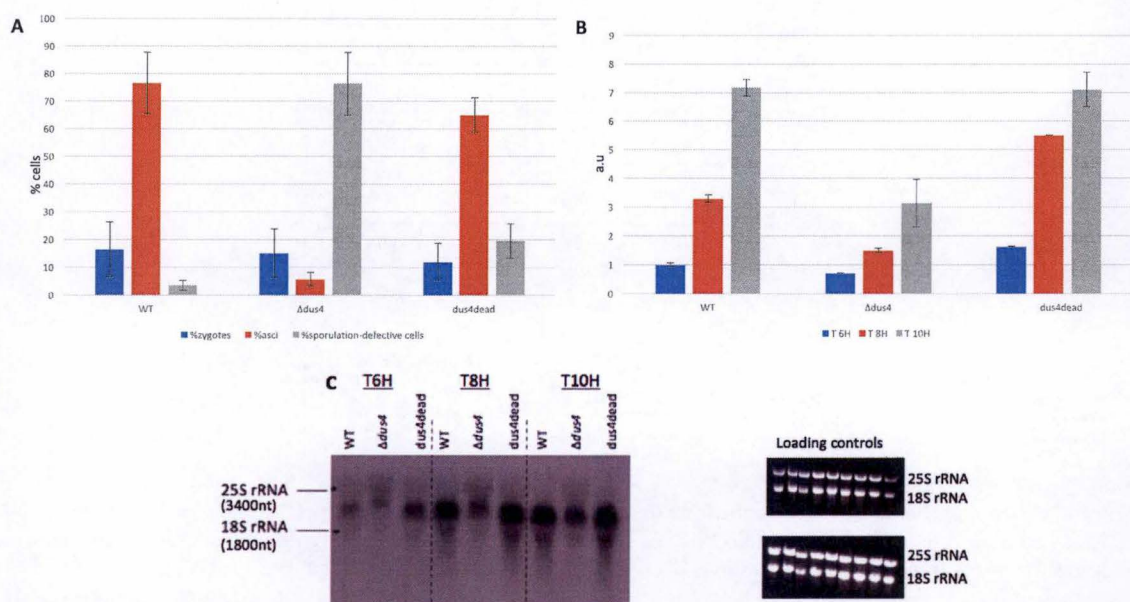


**Figure 17. Exemple théorique de la technique de primer extension.** 1) ARN hypothétique issu d'un extrait total d'ARN et possédant une dihydrouridine (D). 2) Marquage spécifique par lien covalent entre la dihydrouridine et le fluorophore rhodamine (D-R) (échantillons R+). Dans la condition contrôle (échantillons R-), la dihydrouridine n'est pas marquée. 3) Hybridation d'un primer rendu radioactif 4) Réaction de rétro-transcription : en présence de D liée par la rhodamine, la reverse transcriptase est bloquée et le cDNA obtenu est plus court que dans les échantillons dans lesquels la D n'est pas modifiée (échantillons R-) ou non présente. 5) Electrophorèse des cDNA radioactifs produits sur gel polyacrylamide. La bande la plus légère correspond à l'excès de primers, la bande de 40 nucléotides de long correspond au polynucléotide au bout duquel la reverse transcriptase a été bloquée par la D due à la présence de rhodamine. La bande de 80 nucléotides de long correspond au cDNA entier pour lequel l'enzyme a rétro-transcrit jusqu'à l'extrémité 5' (Figure adaptée d'Olivier Finet).



**Figure 18. Primer extension réalisé sur l'ARNt TyrGTA de *Schizosaccharomyces pombe* avec les souches wild type,  $\Delta dus4$ ,  $dus4_{DEAD}$ ,  $\Delta dus1-2-3$ ,  $\Delta dus1-2-3-dus4_{DEAD}$  et  $\Delta dus1-2-3-4$ .** Marquage par rhodamine de différents échantillons totaux d'ARN. R+ correspond à l'échantillon dans lequel la rhodamine lie la dihydrouridine à différentes positions et R- correspond à l'échantillon dans lequel la rhodamine ne lie pas la dihydrouridine. Le primer utilisé mesure 20 nucléotides (flèche 5). La réaction de primer extension a été réalisée et les échantillons ont été chargés sur un gel de polyacrylamide. Les marques correspondent aux segments radioactifs de différentes tailles. **Flèche 1** : stop à 62 nt correspondant à l'extrémité 5', **Flèche 2** : stop à 53 nt correspondant à la modification  $m^1G$ , une autre modification post-transcriptionnelle qui bloque naturellement la réaction de rétro-transcription. **Flèche 3** : stop à 41 nt, correspondant à la dihydrouridine Dus4-dépendante qui est seulement présente dans la souche WT R+ et  $\Delta dus1-2-3$  R+ (position D20a). **Flèche 4** : bande aspécifique, car également présente dans l'échantillon « primer only » (eau à la place d'ARN).





**Figure 19. L'absence de l'activité catalytique de la protéine Dus4 n'est pas responsable du défaut de sporulation observé dans la souche  $\Delta dus4$ .** A) Pourcentage de zygotes (bleu), asques (orange), asques présentant un défaut de sporulation (gris) 48h après l'induction de la sporulation dans les souches wild type (WT),  $\Delta dus4$  et  $dus4^{DEAD}$ . Les barres d'erreur représentent l'écart-type pour trois réplicas biologiques indépendants (représentant 600 cellules). B) qRT-PCR du gène *spo6* 6h, 8h et 10h après l'induction de la sporulation dans les souches wild type (WT),  $\Delta dus4$  et  $dus4^{DEAD}$ . Les barres d'erreur représentent l'écart-type pour deux réplicas techniques, l'ARNm du *house keeping gene act1* a été utilisé pour la normalisation. a.u = arbitrary units (WT 6h = 1) C) Northern Blot du gène *spo6* 6h, 8h and 10h après l'induction de la sporulation dans les souches wild type (WT),  $\Delta dus4$  et  $dus4^{DEAD}$ . Les ARN ribosomiques ont été utilisés comme control de charge, haut : gel d'agarose, bas : membrane de nylon.



Afin de vérifier que la souche créée *dus4*<sub>DEAD</sub> contenait bien une version inactive de la protéine Dus4, un primer extension a été réalisé. Le but de cette technique est d'identifier une modification de l'ARN. Dans le cas précis, elle devait nous permettre de tester l'activité de la protéine Dus4 en détectant la présence ou non d'une dihydrouridine correspondante. Cette méthode est basée sur une rétro-transcription avec un primer spécifique radioactif d'une région désirée. Grâce à un ajout spécifique de rhodamine sur la dihydrouridine, ce qui la rend incapable de s'apparier avec une autre base, la rétro-transcription est arrêtée quand l'enzyme rencontre une dihydrouridine rhodaminisée. La rhodamine, qui est un fluorophore, s'avère nécessaire car la dihydrouridine seule peut s'apparier avec l'adénine et est donc incapable de bloquer le processus de rétro-transcription. Les fragments radioactifs de différentes tailles obtenus sont ensuite séparés par électrophorèse et la présence de dihydrouridine ou non est déduite en fonction de la longueur des fragments obtenus sous différentes conditions (**Fig. 17**).

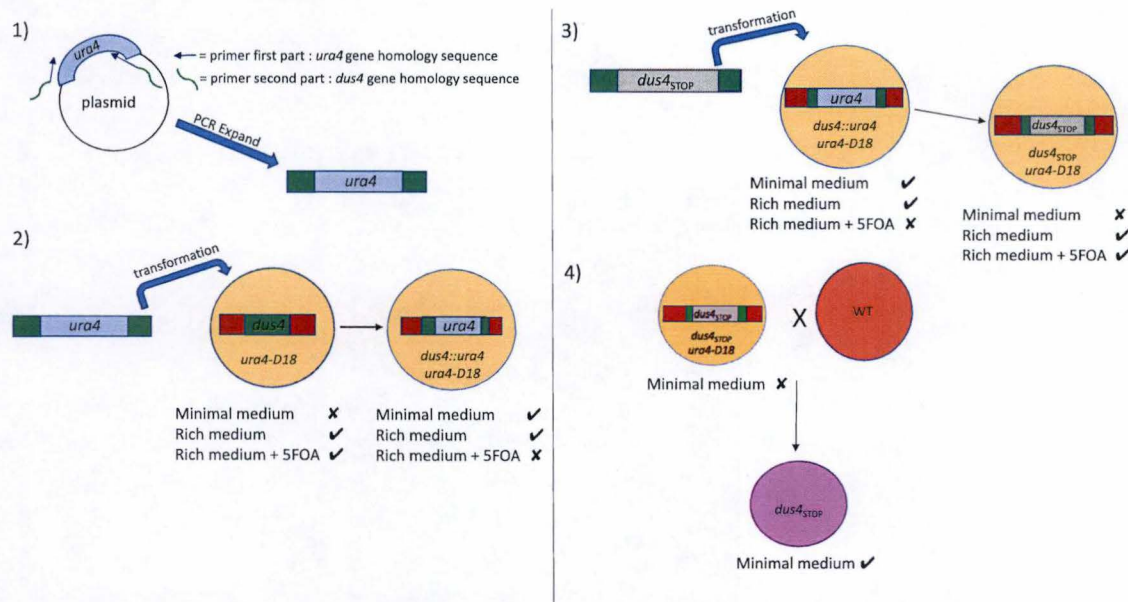
Le primer extension a ici été réalisé à l'aide d'un primer spécifique à un ARN de transfert pour lequel les différentes positions de dihydrouridine sont connues chez *S. pombe*<sup>45</sup>. Les enzymes correspondantes ont ensuite été déduites grâce aux données connues et répertoriées chez *Saccharomyces cerevisiae*<sup>18</sup>.

La figure 18 (**Fig. 18**) montre un primer extension réalisé avec les souches wild type,  $\Delta$ *dus4*, *dus4*<sub>DEAD</sub>,  $\Delta$ *dus1-2-3-dus4*<sub>DEAD</sub> (obtenue par croisement avec une souche  $\Delta$ *dus1-2-3-4*) et les souches préalablement construites  $\Delta$ *dus1-2-3* et  $\Delta$ *dus1-2-3-4*. La construction  $\Delta$ *dus1-2-3-dus4*<sub>DEAD</sub> s'avérait être un contrôle supplémentaire pour confirmer que l'absence d'activité catalytique dans la souche *dus4*<sub>DEAD</sub> était bien due à la mutation du site catalytique du gène *dus4* uniquement et pas à une possible intégration du gène *dus4*<sub>DEAD</sub> dans les loci des autres *Dus*. La dihydrouridine correspondant à l'action de la dihydrouridine synthase 4 est la dihydrouridine en position D20a et le fragment de cDNA correspondant fait 41 nucléotides de long. Cette modification apparaît présente dans les souches wild type et  $\Delta$ *dus1-2-3* mais pas dans les souches construites  $\Delta$ *dus4*, *dus4*<sub>DEAD</sub> et  $\Delta$ *dus1-2-3-dus4*<sub>DEAD</sub>. Ces résultats confirment que la souche *dus4*<sub>DEAD</sub> est bien catalytiquement inactive pour la protéine Dus4, comme attendu.

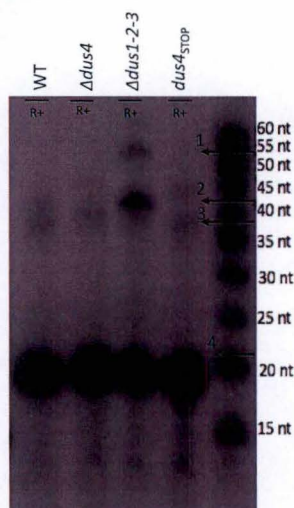
Ensuite, le pourcentage de cellules présentant un défaut de sporulation dans la souche *dus4*<sub>DEAD</sub> a été mesuré. Comme indiqué dans la figure 19 (**Fig. 19A**), la souche *dus4*<sub>DEAD</sub> présente un phénotype similaire à la souche wild type, avec une majorité d'asques possédant des spores sans défaut apparent (65% et 76%, respectivement), bien que jusqu'à 20 % de sporulation defective cells aient quand même été répertoriés.

Pour caractériser le phénotype de défaut de sporulation au niveau moléculaire, l'expression du gène *spo6* a été analysée par qRT-PCR (**Fig. 19B**) et par Northern Blot (**Fig. 19C**), à différents temps après l'induction de la sporulation. Il s'avère que l'expression du gène *spo6* dans la souche *dus4*<sub>DEAD</sub> est comparable à son expression dans la souche wild type, rejoignant l'observation précédente réalisée par comptages. Pris ensemble, ces résultats appuient le fait que l'activité catalytique de la protéine Dus4 ne semble pas nécessaire à une sporulation correcte chez *S. pombe*, et que ce n'est donc pas son absence dans la souche  $\Delta$ *dus4* qui explique, en tous cas totalement, le défaut de sporulation observé.



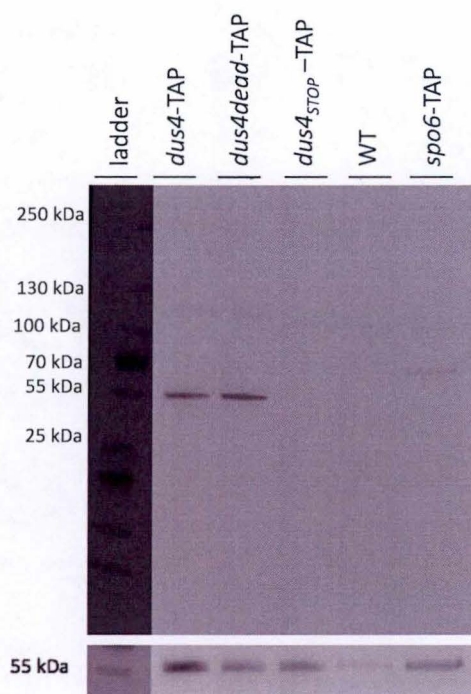


**Figure 20. Construction de la souche *dus4*<sub>STOP</sub>.** 1) PCR Expand (PCR à haute-fidélité répliquative) d'un plasmide porteur du gène *ura4* avec des primers conçus contenant une séquence d'homologie avec le gène *ura4* et le gène *dus4*. 2) Transformation du produit PCR dans une souche porteuse du marqueur d'auxotrophie *ura4-D18* (donc incapable de pousser sur milieu pauvre). La souche obtenue *dus4::ura4 ura4-D18* a été sélectionnée par sa capacité à pousser sur milieu pauvre mais pas sur milieu riche + 5-FOA (métabolite toxique sur les souches *ura+*). 3) Un fragment de gène (*dus4*<sub>STOP</sub> en gris) correspondant à la portion du gène *dus4* dans laquelle huit codons STOP ont été introduits a été transformé dans la souche *dus4::ura4 ura4-D18* montrée à l'étape 2. La souche obtenue *dus4STOP ura4-D18* était capable de pousser sur milieu riche + 5-FOA mais pas sur milieu pauvre. 4) Un croisement entre la souche *dus4STOP ura4-D18* et une souche wild type a été réalisé dans le but de perdre le marqueur d'auxotrophie *ura4-D18*. Tous les clones obtenus ont été vérifiés par PCR aux différentes étapes de la construction et la souche finale *dus4STOP* a été vérifiée par séquençage.

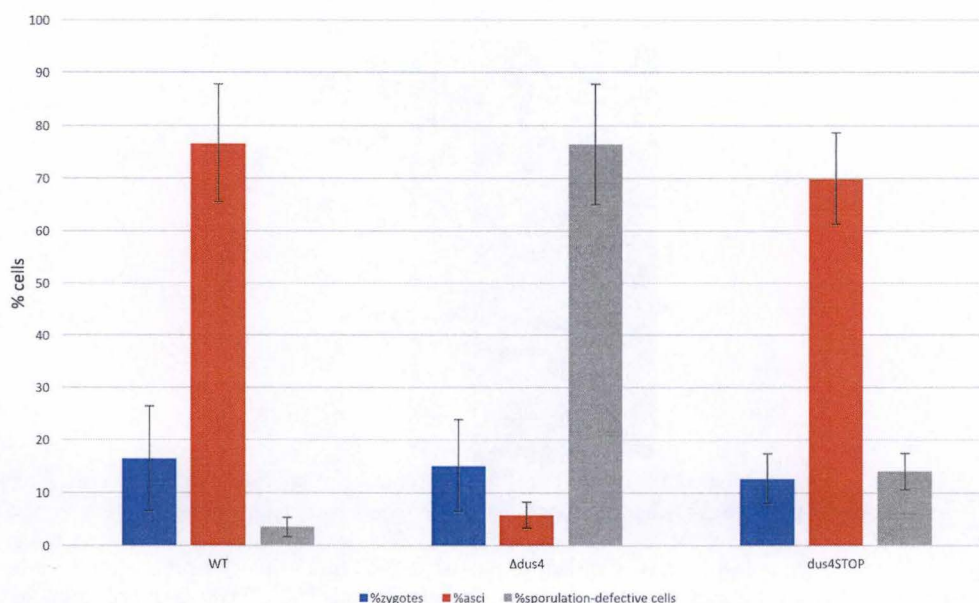


**Figure 21. Primer extension réalisé sur l'ARNt TyrGTA de Schizosaccharomyces pombe avec les souches wild type,  $\Delta$ *dus4*,  $\Delta$ *dus1-2-3* et *dus4*<sub>STOP</sub>.** Marquage par rhodamine de différents échantillons totaux d'ARN. R+ correspond à l'échantillon dans lequel la rhodamine lie la dihydrouridine à différentes positions et R- correspond à l'échantillon dans lequel la rhodamine ne lie pas la dihydrouridine. Le primer utilisé mesure 20 nucléotides de long (flèche 4) et marqué radioactivement avec une polynucléotide kinase. La réaction de primer extension a été réalisée et les échantillons ont été chargés sur un gel de polyacrylamide. Les marques correspondent aux segments radioactifs de différentes tailles. **Flèche 1** : stop à 53nt, correspondant à la modification m<sup>1</sup>G, une autre modification post-transcriptionnelle qui bloque naturellement la réaction de transcription inverse. **Flèche 2** : stop à 41 nt correspondant à la dihydrouridine en position D20a, après l'intervention de la protéine Dus4. **Flèche 3** : bande aspécifique.





**Figure 22. Western Blot des souches wild type, *dus4-TAP*, *dus4<sup>DEAD</sup>-TAP*, *dus4<sup>STOP</sup>-TAP*.** Haut: Détection des protéines TAP-tagées avec un anticorps anti-TAP (peroxidase anti-peroxidase). La révélation a été réalisée par chimioluminescence. La taille attendue de la protéine Dus4-TAP est de 58kDa. La souche *spo6-TAP* a été utilisée comme contrôle avec une taille attendue de 76kDa. Bas : Contrôle de charge par le marquage de la tubuline avec un anticorps secondaire dans les souches wild type, *dus4-TAP*, *dus4<sup>DEAD</sup>-TAP*, *dus4<sup>STOP</sup>-TAP*, et *spo6-TAP*. La révélation a été réalisée avec le réactif « Enhanced Chemiluminescence Lightning » et le logiciel ImageQuant. La taille attendue de la tubuline est de 55kDa.



**Figure 23. L'absence de la protéine Dus4 n'est pas responsable du phénotype de défaut de sporulation observé.** Pourcentage de zygotes (bleu), asques (orange), asques présentant un défaut de sporulation (gris) 48h après l'induction de la sporulation dans les souches wild type (WT),  $\Delta dus4$  et *dus4<sup>STOP</sup>*. Les barres d'erreur représentent l'écart-type pour trois répliques biologiques indépendants (représentant 600 cellules).



#### IV.1.3 Etude de l'implication de la protéine Dus4 dans le défaut de sporulation observé dans la souche $\Delta dus4$

Bien que l'absence d'activité catalytique ne semble pas jouer de rôle dans le phénotype de défaut de sporulation observé dans la souche  $\Delta dus4$ , l'implication de la protéine indépendamment de son activité ne peut pas être exclue par l'analyse de la seule souche  $dus4_{DEAD}$ . C'est pourquoi une souche incapable de produire la protéine Dus4 a été créée. La différence essentielle avec un mutant de délétion est que le locus est toujours présent et intact à l'exception des mutations ponctuelles décrites ci-dessous.

La nouvelle souche, appelée  $dus4_{STOP}$  a été obtenue par le changement de 8 codons en STOP dans la région du gène *dus4* codant pour la protéine Dus4. Dans cette souche, la traduction du transcrit encore présent devait donc normalement être empêchée alors que la transcription ne devrait pas être affectée. La souche  $dus4_{STOP}$  a été construite en suivant le même schéma de construction que la souche  $dus4_{DEAD}$ , par une stratégie de sélection en deux étapes, avec l'aide d'une souche *dus4::ura4 ura4-D18*, et d'une version synthétique du gène *dus4* possédant ici les 8 codons STOP (**Fig. 20**).

L'absence de dihydrouridine en position D20a, prédite sur base de l'absence de la protéine Dus4, a d'abord été vérifiée par un primer extension réalisé sur les souches wild type,  $\Delta dus4$ ,  $\Delta dus1-2-3$  et  $dus4_{STOP}$  (**Fig. 21**). Comme attendu, la modification D20a n'est pas détectée dans la souche  $dus4_{STOP}$ .

Ensuite, les souches *dus4* WT,  $dus4_{DEAD}$ ,  $dus4_{STOP}$  TAP-tagées[1] ont été créées. Les versions taggées des protéines Dus4, Dus4<sub>DEAD</sub> et Dus4<sub>STOP</sub> ont été générées et analysées par Western Blot afin de vérifier la présence ou l'absence de la protéine Dus4 (**Fig. 22**). La protéine Dus4 était présente dans les souches *dus4-TAP* et  $dus4_{DEAD-TAP}$  mais pas dans la souche  $dus4_{STOP-TAP}$ , confirmant l'authenticité des souches que nous avons construites.

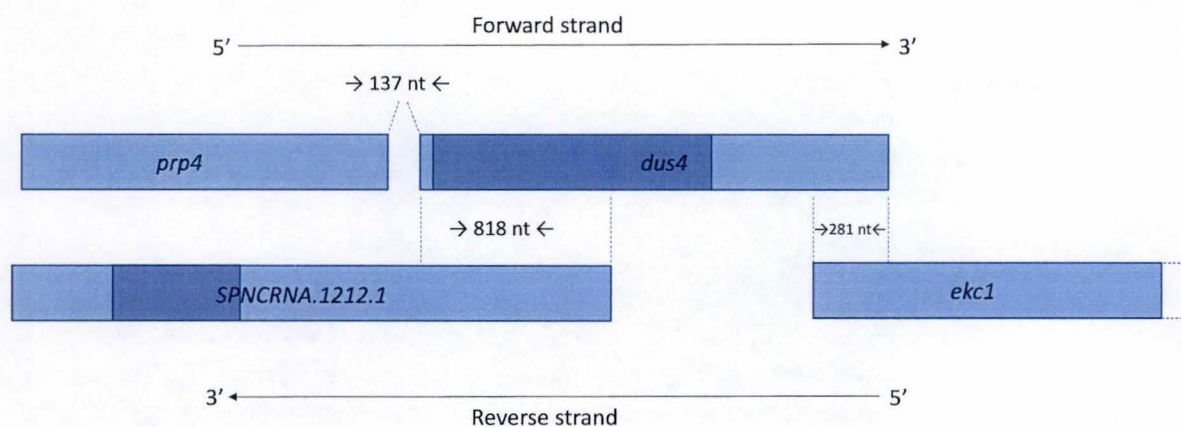
Une analyse phénotypique a été réalisée dans la souche  $dus4_{STOP}$  en comparaison avec les souches wild type et  $\Delta dus4$ . Comme indiqué dans la figure 23 (**Fig. 23**), la souche  $dus4_{STOP}$  sporule d'une manière comparable à la souche wild type (69,9% et 76,6% d'asques bien formés respectivement). Ceci exclut un rôle structural de la protéine Dus4 dans le processus de sporulation et suggère que c'est l'altération du locus d'ADN dans la souche  $\Delta dus4$  qui est causale pour le défaut de sporulation observé.

#### IV.1.4 Etude de l'implication du transcrit *dus4* dans le défaut de sporulation observé dans la souche $\Delta dus4$

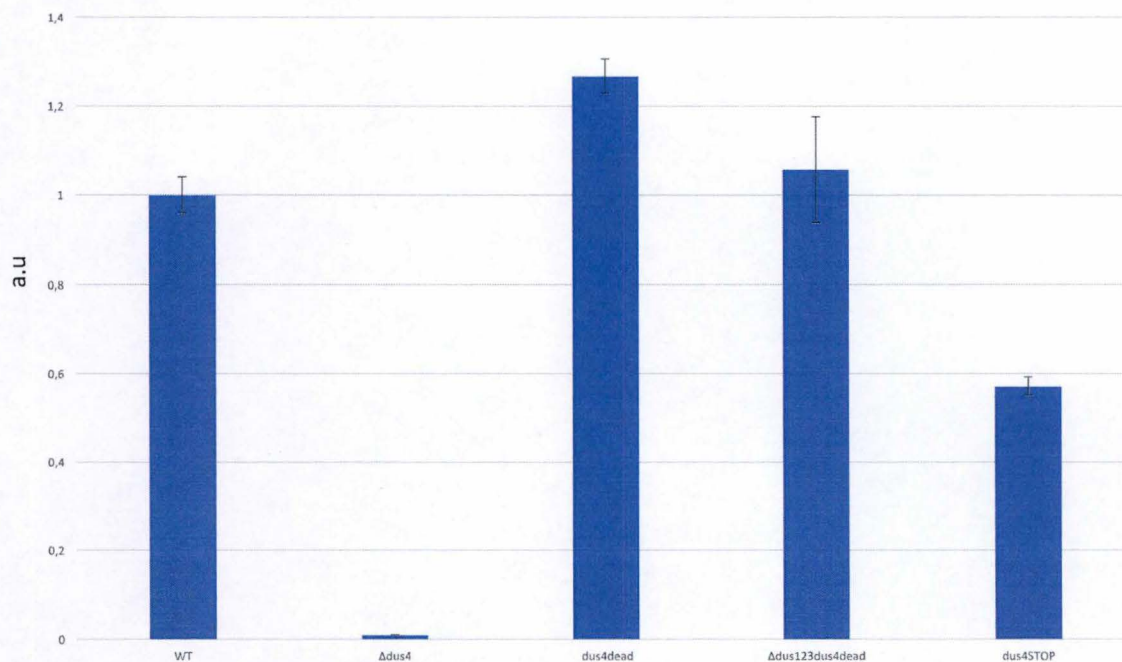
A ce stade des résultats, il pouvait être affirmé que le défaut de sporulation observé dans la souche  $\Delta dus4$  n'était pas dû à l'absence de la protéine provoquée par la délétion du gène. Ceci concernant aussi bien la perte catalytique que la perte totale de la protéine.

L'implication de la perte d'un transcrit émanant du locus *dus4* dans le défaut de sporulation observé a ensuite été envisagée. Un moyen de vérifier cette hypothèse était de savoir si un ARN<sup>[UB2]</sup> correspondant à ce locus *dus4* était encore présent dans la souche  $dus4_{STOP}$  qui, pour rappel, sporule d'une façon similaire à la souche wild type.



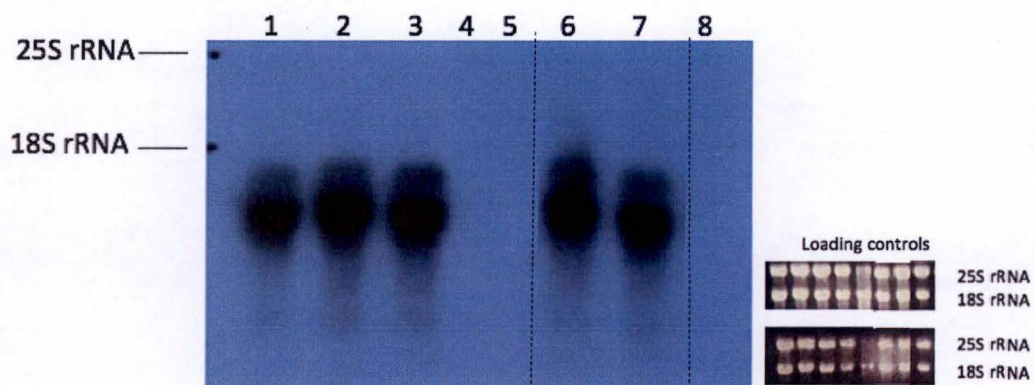


**Figure 24. Schématisation de l'environnement génomique du gène *dus4*.** Le gène *dus4* (1656pb) est situé sur le brin sens et est uniquement espacé du gène *prp4* (1978 pb) par 137 nucléotides. Face au gène *dus4* sur le brin anti-sens se trouvent le gène codant pour le long ARN non codant *SPNCRNA.1212.1* (qui possède un chevauchement de 818 nt avec le gène *dus4*, 2905 pb) et le gène codant pour la protéine Ekc1 (qui possède un chevauchement de 281 nt avec le gène *dus4*, 3451pb) (données obtenues sur le site [www.pombase.org](http://www.pombase.org)). Les régions surlignées en bleu plus foncé correspondent aux sondes utilisées dans la réalisation des Northern Blot brin non spécifique *dus4* et *prp4*. La sonde utilisée pour le Northern Blot *dus4* fait 748 nt et est située dans le seul exon de *dus4*. Tandis que la sonde utilisée pour le Northern Blot *prp4* fait 343 nt et est située dans l'exon 1 du gène.



**Figure 25. Niveau d'expression de l'ARN produit à partir du locus *dus4* dans les souches wild type (WT),  $\Delta dus4$ , *dus4*<sub>DEAD</sub>,  $\Delta dus123dus4$ <sub>DEAD</sub> et *dus4*<sub>STOP</sub>.** Les barres d'erreur représentent la SEM pour trois répliques techniques, l'ARNm du house keeping gene *act1* a été utilisé pour la normalisation. a.u = arbitrary units (WT = 1)





**Figure 26. Northern blot du gène *dus4*.** 1: WT 2:  $\Delta dus2$  3:  $\Delta dus123$  4:  $\Delta dus4$  (*dus4::ble<sup>R</sup>*) 5:  $\Delta dus4$  (*dus4::ble<sup>R</sup> leu-*) 6:  $\Delta dus1-2-3$  *dus4<sub>DEAD</sub>* 7: *dus4<sub>STOP</sub>* 8:  $\Delta dus1-2-3-4$ . Les ARN ribosomiques ont été utilisés comme contrôle de charge, haut : gel d'agarose, bas : membrane de nylon. *dus4* : 1666nt, 18S : 1800 n, 25S : 3400 nt



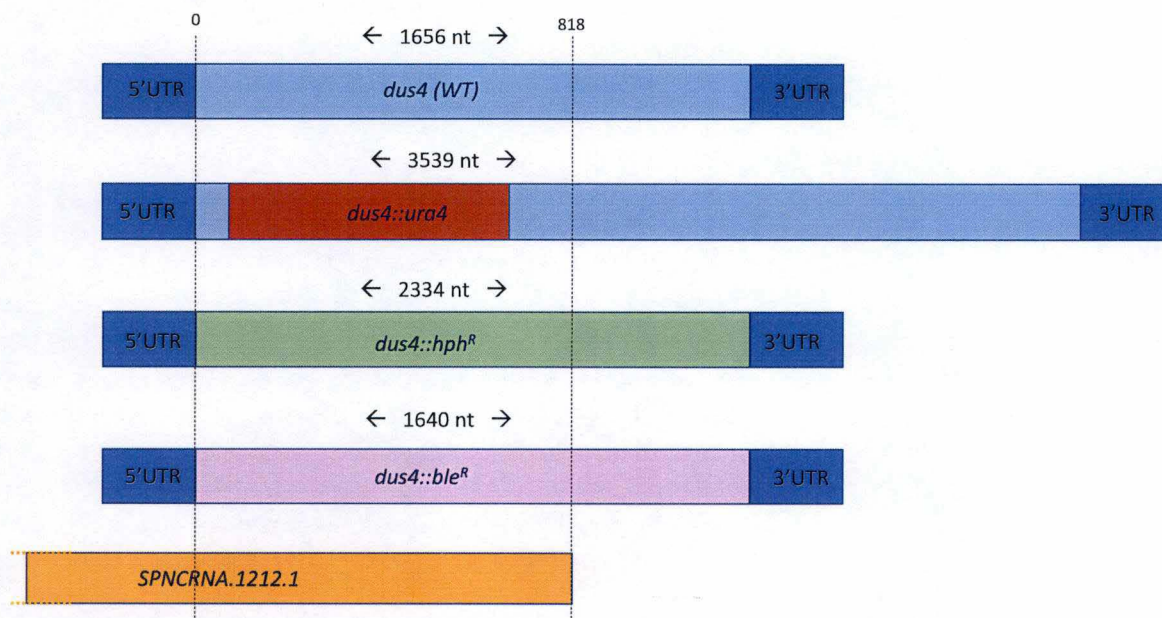
Si la perte du transcrit dans cette souche ne provoque pas de défaut de sporulation (voir résultats précédents), alors il peut être extrapolé que la perte du transcrit dans la souche  $\Delta dus4$  ne provoque pas non plus le défaut de sporulation observé. A contrario, si le transcrit *dus4* est encore présent dans la souche *dus4*<sub>STOP</sub>, alors il peut potentiellement être envisagé que son absence dans la souche  $\Delta dus4$  puisse jouer un rôle dans le défaut de sporulation observé, puisque le fait de ne pas le perdre dans la souche *dus4*<sub>STOP</sub> ne provoque pas le défaut de sporulation.

Un premier moyen de vérifier cette hypothèse était de quantifier l'ARN messenger par qRT-PCR. Etant donné que la quantification de l'ARN par qRT-PCR se fait sur l'ARN total après son extraction et purification, il était important de quantifier uniquement le transcrit *dus4* et non le transcrit qui lui fait face sur le brin anti-sens, c'est-à-dire ici le long ARN non codant *SPNCRNA.1212.1*. Des primers spécifiques de la région du transcrit *dus4* uniquement (sans chevauchement avec le long ARN non codant *SPNCRNA.1212.1*) ont donc été conçus et utilisés. L'environnement génomique du gène *dus4* est schématisé dans la figure 24 (**Fig. 24**). La qRT-PCR a été réalisée sur les souches WT,  $\Delta dus4$ , les souches construites *dus4*<sub>DEAD</sub>,  $\Delta dus1-2-3-dus4$ <sub>DEAD</sub> et la souche *dus4*<sub>STOP</sub>. Le transcrit *dus4*, normalisé à 1 dans la souche wild type par rapport au *housekeeping gene act1*, n'apparaît pas dans la souche  $\Delta dus4$  (0,009), et est quantifié à un niveau d'expression similaire à la souche wild type dans les souches *dus4*<sub>DEAD</sub> (1,27) et  $\Delta dus1-2-3 dus4<sub>DEAD</sub> (1,05) (**Fig. 25**). Concernant la souche *dus4*<sub>STOP</sub>, le transcrit *dus4* est toujours présent, bien qu'à un niveau d'expression diminué comparativement à la souche wild type (0,57).$

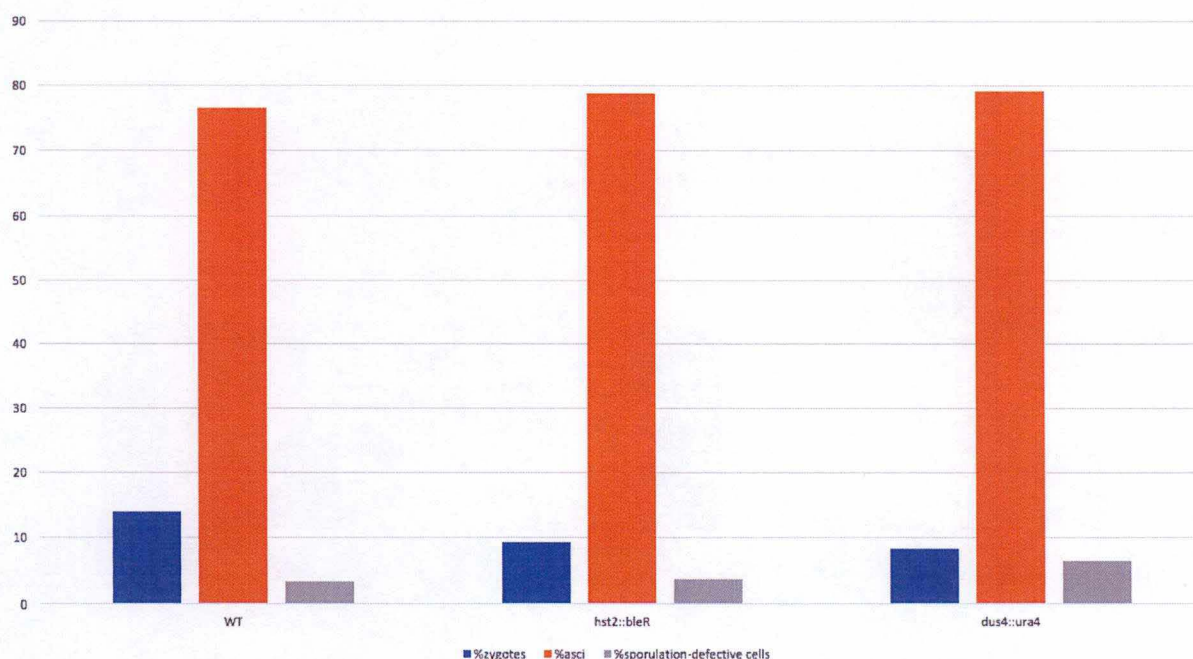
Pour appuyer ce résultat et ici visualiser qualitativement la présence de l'ARN messenger de *dus4* dans la souche construite *dus4*<sub>STOP</sub>, un Northern Blot a été réalisé. Suite à de nombreux essais infructueux de réalisation d'un Northern Blot brin spécifique, toujours dans l'objectif de quantifier uniquement l'ARN de *dus4* et non le long ARN non codant *SPNCRNA.1212.1*, un Northern Blot brin non spécifique a finalement été réalisé (**Fig. 26**). L'échec de la réalisation du Northern Blot brin spécifique peut par exemple s'expliquer par un défaut de construction de la sonde utilisée ou par une moindre efficacité de la radioactivité, étant donné qu'aucun marquage n'était visible à l'œil nu. La position de la sonde utilisée est représentée en figure 24 (**Fig. 24**). Les souches contrôle WT,  $\Delta dus4$  (*dus4::ble<sup>R</sup>*),  $\Delta dus4$  (*dus4::ble<sup>R</sup> leu-*),  $\Delta dus2$ ,  $\Delta dus1-2-3$ ,  $\Delta dus1-2-3-4$  ont été utilisées en plus de la souche *dus4*<sub>STOP</sub>. Deux bandes sont clairement distinguables. Etant donné la non spécificité de la sonde radioactive utilisée, il peut être supposé que la deuxième bande corresponde au gène faisant face au gène *dus4* sur le brin anti-sens, ici : le long ARN non codant *SPNCRNA.1212.1*. Concernant la souche *dus4*<sub>STOP</sub> (puits numéro 7), la bande correspondant à l'ARN messenger de *dus4* est clairement visible. Contrairement à ce qui est quantifié dans la qRT-PCR, aucune différence d'expression notable comparativement à la souche WT ne peut être ici mise en avant.

Il peut donc être émis comme hypothèse que la perte du transcrit *dus4* dans la souche  $\Delta dus4$  joue un rôle dans le défaut de sporulation observé, étant donné la présence incontestable du transcrit *dus4* dans la souche *dus4*<sub>STOP</sub>.



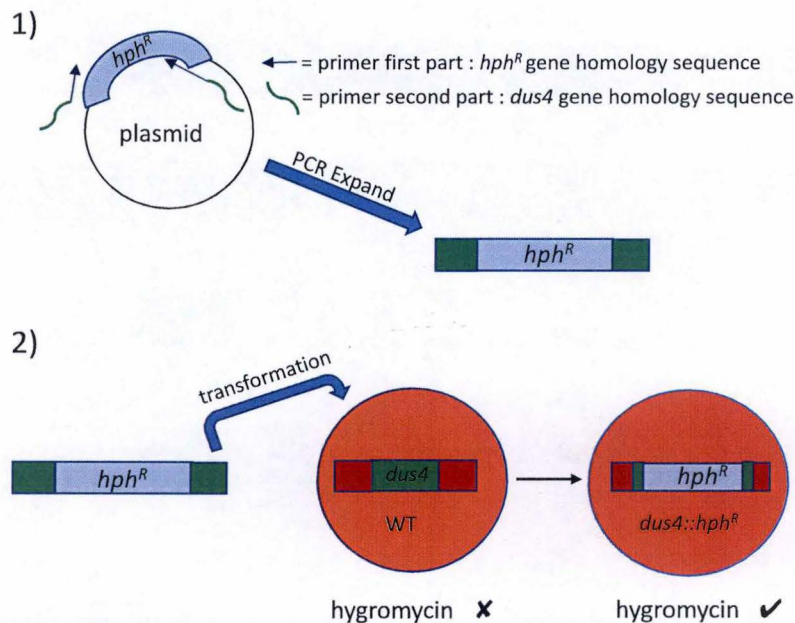


**Figure 27. Représentation des portions du gène *dus4* remplacées dans les différentes souches utilisées pour analyser l'impact de la cassette de résistance *ble<sup>R</sup>*. La CDS est complètement remplacée dans les souches *dus4::ble<sup>R</sup>* et *dus4::hph<sup>R</sup>*, mais pas dans la souche *dus4::ura4*. Le long ARN non codant *SPNCRNA.1212.1* situé sur le brin complémentaire de *dus4* chevauche *dus4* sur 818 nucléotide.**

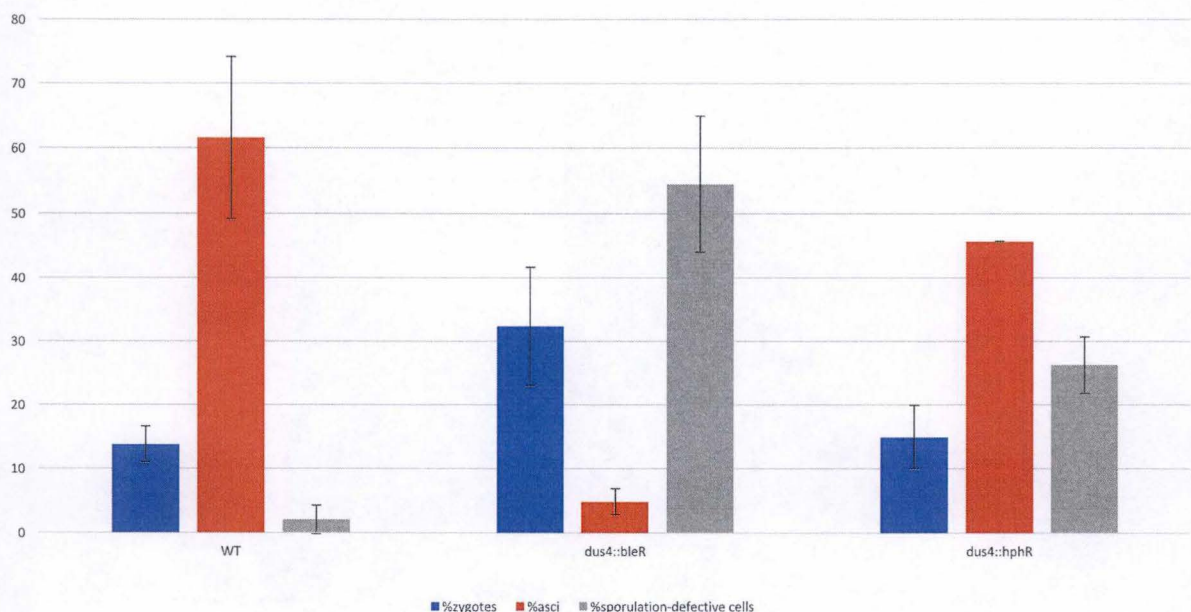


**Figure 28. La cassette de résistance *ble<sup>R</sup>* seule et le remplacement de la majorité de la CDS de *dus4* sont insuffisants pour provoquer le défaut de sporulation. A) Pourcentage de zygotes (bleu), d'asques (orange), d'asques avec un défaut de sporulation (gris) 48h après l'induction de la sporulation dans les souches wild type (WT), *hst2::ble<sup>R</sup>* et *dus4::ura4*. Les barres d'erreur représentent l'écart-type pour trois répliques biologiques indépendantes (représentant 600 cellules).**





**Figure 29. Construction de la souche *dus4::hph<sup>R</sup>*** 1) PCR Expand (PCR à haute-fidélité répliquative) d'un plasmide porteur du gène codant pour la protéine de résistance *hph<sup>R</sup>* avec des primers conçus qui contiennent une homologie de séquence avec le gène *hph<sup>R</sup>* et le gène *dus4*. 2) Transformation du produit PCR dans une souche wild type. Les clones *dus4::hph<sup>R</sup>* obtenus ont été sélectionnés sur un milieu de sélection (hygromycine) et vérifiés par PCR.



**Figure 30. La cassette de résistance *hph<sup>R</sup>* provoque un phénotype de sporulation intermédiaire.** A) Pourcentage de zygotes (bleu), asques (orange), asques présentant un défaut de sporulation (gris) 48h après l'induction de la sporulation dans les souches wild type (WT),  $\Delta dus4$  (*dus4::ble<sup>R</sup>*) et *dus4::hph<sup>R</sup>*. Les barres d'erreur représentent l'écart-type pour trois répliques biologiques indépendants (représentant 600 cellules).



#### IV.1.5 Etude de l'implication de la cassette de résistance utilisée dans la souche $\Delta dus4$ dans le défaut de sporulation observé

Les données précédentes suggèrent que c'est le fragment de chromatine localisé au locus *dus4* ou un transcrit émanant de cette région qui est requis pour une sporulation efficace. Nous avons aussi voulu vérifier que la cassette de résistance employée, *ble<sup>R</sup>*, ne jouait pas de rôle dans le défaut de sporulation observé dans la souche  $\Delta dus4$ , indépendamment de sa localisation génomique. Dans la souche  $\Delta dus4$ , le gène codant pour la protéine de résistance à la bléomycine remplace le gène *dus4* après transformation d'une souche wild type, depuis le codon START et jusqu'au codon STOP, et ce grâce à la spécificité des primers conçus et utilisés (Fig. 27).

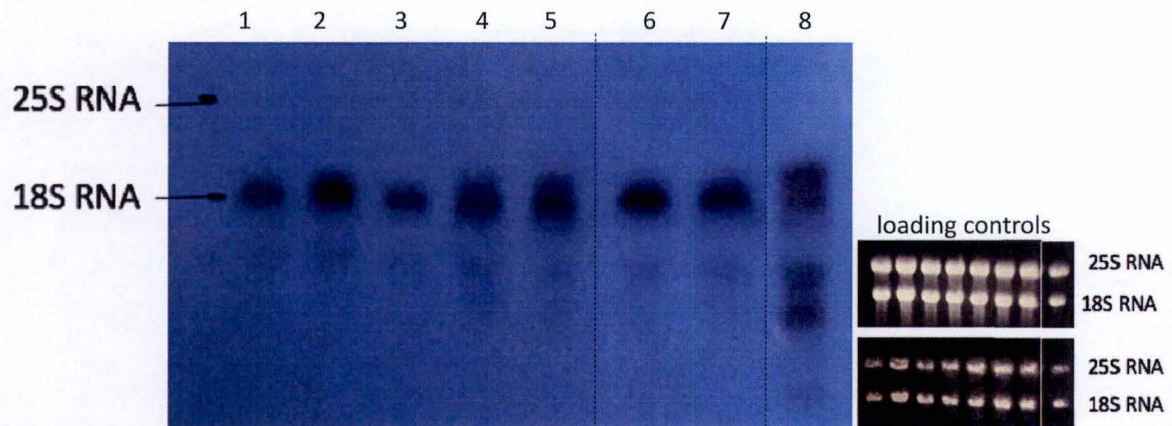
Premièrement, après avoir provoqué la sporulation, des comptages ont été effectués dans une souche possédant un autre gène remplacé par la cassette de résistance *ble<sup>R</sup>*, le gène *hst2*, codant pour une histone déacétylase (impliquée notamment dans la répression de l'expression des gènes subtélomériques et dans le silençage de l'ADN ribosomique <sup>46</sup>). Par comptages cellulaires, la souche *hst2::ble<sup>R</sup>* ne présentait aucun défaut de sporulation (3,66 % d'asques défectueux et 3,33 % dans la souche WT) (Fig. 28), excluant l'impact isolé de la cassette de résistance sur le défaut de sporulation observé dans la souche  $\Delta dus4$ .

Deuxièmement, des comptages ont été effectués sur une souche dans laquelle le gène *dus4* est partiellement délété par un autre gène, le gène *ura4*. Cependant, ici, toute la CDS (coding DNA sequence) du gène *dus4* n'est pas remplacée, et seule une petite partie du long ARN non codant *SPNCRNA.1212.1* est remplacée (Fig. 27). Par comptages cellulaires et après avoir engendré la sporulation, la souche *dus4::ura4* ne présentait pas de défaut de sporulation (6,39 % d'asques défectueux et 3,33 % dans la souche WT) (Fig. 28), excluant l'impact de la perte de la partie de la CDS de *dus4* remplacée par *ura4* dans le défaut de sporulation observé dans la souche  $\Delta dus4$ .

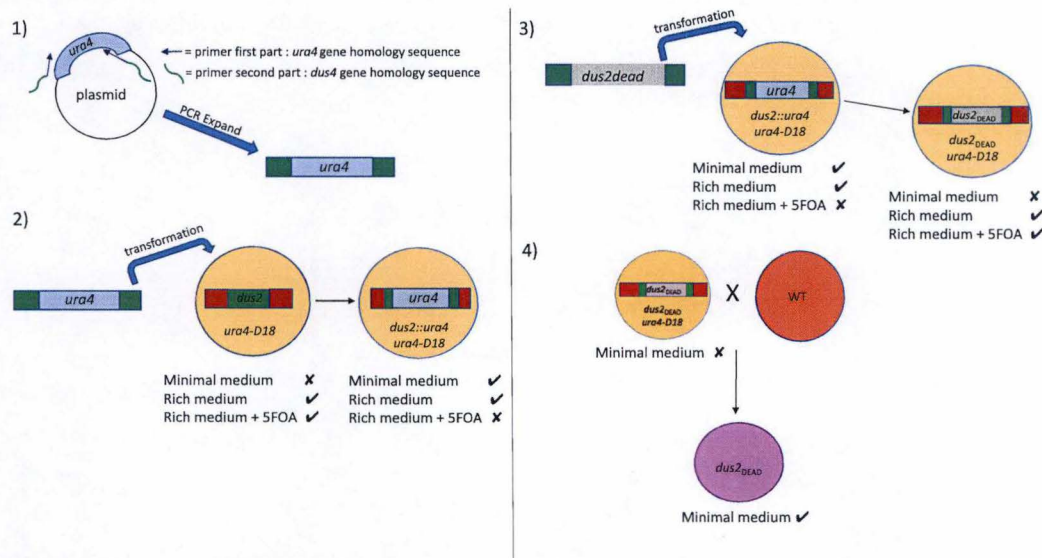
Troisièmement, un autre gène codant pour une autre cassette de résistance a été utilisé pour construire une seconde souche où le gène *dus4* est remplacé sur toute la CDS. Il s'agit du gène *hph<sup>R</sup>*, conférant la résistance à l'antibiotique hygromycine, qui a aussi ici été introduit par transformation d'une souche wild type. La construction réalisée suit les mêmes étapes que pour la construction de la souche  $\Delta dus4_{BIS}$  (Fig. 29). Dans cette souche, l'absence du transcrit *dus4* a été confirmée par la réalisation d'un Northern Blot (voir Fig. 26). Des comptages cellulaires ont ensuite été réalisés sur la souche *dus4::hph<sup>R</sup>* construite. Il s'avère que cette souche présente un défaut de sporulation mais avec un pourcentage intermédiaire : 30,8 % d'asques défectueux (Fig. 30).

En regard de ces résultats, il peut être affirmé que la cassette de résistance *ble<sup>R</sup>* seule n'est pas responsable du défaut de sporulation observé. Mais il ne peut pas être affirmé qu'elle ne joue pas de rôle du tout étant donné qu'une grande partie du gène *dus4* remplacé par un autre gène ne présente pas de défaut de sporulation et qu'une autre cassette de résistance ne provoque pas un aussi grand défaut de sporulation.



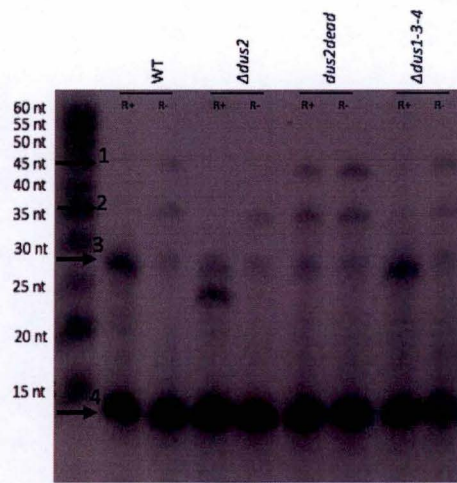


**Fig 31. Northern Blot du gène *prp4*.** 1: WT 2:  $\Delta dus2$  3:  $\Delta dus1-2-3$  4:  $\Delta dus4$  (*dus4::ble<sup>R</sup>*) 5:  $\Delta dus4$  (*dus4::ble<sup>R</sup> leu-*) 6:  $\Delta dus1-2-3$  *dus4<sub>DEAD</sub>* 7: *dus4<sub>STOP</sub>* 8:  $\Delta dus1-2-3-4$ . Les ARN ribosomiques ont été utilisés comme contrôle de charge, haut : gel d'agarose, bas : membrane de nylon. *prp4* : 1900nt, 18S : 1800 n, 25S : 3400 nt.

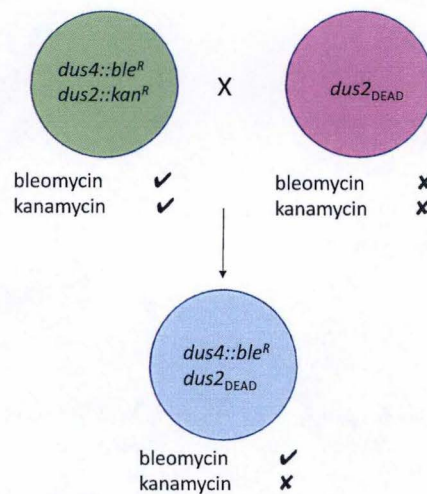


**Figure 32. Construction de la souche *dus2<sub>DEAD</sub>*.** 1) PCR Expand (PCR à haute fidélité répliquative) d'un plasmide porteur du gène *ura4* avec des primers conçus contenant une séquence d'homologie avec le gène *ura4* et le gène *dus2*. 2) Transformation du produit PCR dans une souche porteuse du marqueur d'auxotrophie *ura4-D18* (donc incapable de pousser sur milieu pauvre). La souche obtenue *dus2::ura4 ura4-D18* a été sélectionnée par sa capacité à pousser sur milieu pauvre mais pas sur milieu riche + 5-FOA (un métabolite toxique sur les souches *ura+*). 3) Un fragment de gène (*dus2<sub>DEAD</sub>* en gris) correspondant à la portion du gène *dus2* dans laquelle deux mutations pour l'activité catalytique de la protéine Dus2 ont été conçues (C108A K149A) a été transformé dans la souche *dus2::ura4 ura4-D18* montrée à l'étape 2. La souche obtenue *dus2<sub>DEAD</sub> ura4-D18* était capable de pousser sur milieu riche + 5-FOA mais pas sur milieu pauvre. 4) Un croisement entre la souche *dus2<sub>DEAD</sub> ura4-D18* et une souche wild type a été réalisé dans le but de perdre le marqueur d'auxotrophie *ura4-D18*. Tous les clones obtenus ont été vérifiés par PCR aux différentes étapes de la construction et la souche finale *dus2<sub>DEAD</sub>* a été vérifiée par séquençage.





**Figure 33. Primer extension réalisé sur l'ARNt Asp06 de *Schizosaccharomyces pombe* avec les souches wild type,  $\Delta dus2$ ,  $\Delta dus2_{DEAD}$  et  $\Delta dus1-3-4$ .** Marquage par rhodamine de différents échantillons totaux d'ARN. R+ correspond à l'échantillon dans lequel la rhodamine lie la dihydrouridine à différentes positions et R- correspond à l'échantillon dans lequel la rhodamine ne lie pas la dihydrouridine. Le primer utilisé mesure 16 nucléotides (**flèche 4**) et est marqué radioactivement avec une polynucléotide kinase. La réaction de primer extension a été réalisée et les échantillons ont été chargés sur un gel de polyacrylamide. Les marques correspondent aux segments rendus radioactifs de différentes tailles. **Flèche 1** : stop à 48nt, correspondant à l'extrémité 5' de l'ARNt. **Flèche 2** : potentielle autre modification post-transcriptionnelle bloquant la reverse transcription. 4. **Flèche 3** : stop à 29 nt, correspondant à la dihydrouridine en position D19 obtenue suite à l'action de la dihydrouridine synthase 2.



**Figure 34. Construction de la souche  $\Delta dus4 \Delta dus2_{DEAD}$ .** Croisement entre une souche  $\Delta dus4 \Delta dus2$  et une souche  $dus2_{DEAD}$  sur un milieu carencé en azote pour induire la sporulation. Les clones  $dus4::ble^R \Delta dus2_{DEAD}$  obtenus ont été sélectionnés sur deux milieux de sélection (bléomycine et kanamycine) et vérifiés par PCR.



#### IV.1.6 Etude de l'implication de l'environnement génomique dans le défaut de sporulation observé dans la souche $\Delta dus4$

Etant donné l'environnement génomique dense du gène *dus4* qui pourrait potentiellement être affecté par l'introduction de la cassette de résistance et l'importance du gène voisin *prp4* (**Fig. 24**) dans la ségrégation chromosomique <sup>47</sup>, un Northern Blot a été réalisé afin de vérifier si l'expression du transcrit de *prp4* était affectée. Le gène *prp4* a ici retenu notre attention car il s'avérait être sous-exprimé dans la souche  $\Delta dus4$  (données RNAseq, non montrées) et être un gène essentiel jouant également un rôle dans le splicing de certains introns (*spo6* semble cependant ne pas être concerné) <sup>48</sup>.

Les souches WT,  $\Delta dus4$  (*dus4::ble<sup>R</sup>*),  $\Delta dus4$  (*dus4::ble<sup>R</sup> leu-*),  $\Delta dus2$ ,  $\Delta dus1-2-3$ ,  $\Delta dus1-2-3-4$ , *dus4<sub>STOP</sub>*, ont été utilisées. La position de la sonde utilisée est représentée en figure 24 (**Fig. 24**). Plusieurs constatations peuvent être faites. Premièrement, le gène *prp4* est visible dans tous les puits, sa taille étant de 1900 nt. Deuxièmement, une bande supplémentaire apparaît dans chaque souche délétée en *dus4* (puits 4, 5 et 8). Et finalement, d'autres bandes de taille plus faible sont également visibles dans tous les puits mais de manière plus prononcée dans le puits 8, correspondant à la souche  $\Delta dus1-2-3-4$ . Etant donné que le Northern Blot réalisé était brin non spécifique, il peut être supposé que les autres bandes marquées puissent correspondre au gène localisé sur le brin complémentaire du brin contenant *prp4*, ici : le long ARN non codant *SPNCRNA.1212.1*. Ce résultat est discuté plus en détails dans la partie IV) discussion (**Fig. 31**).

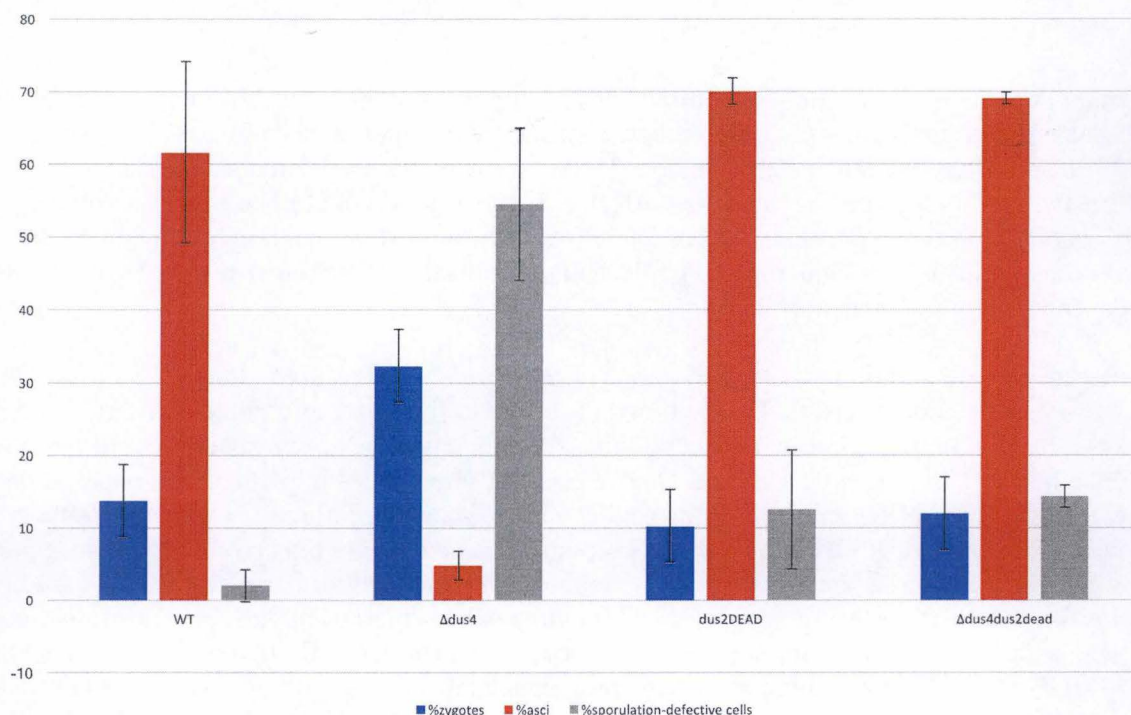
#### IV.1.7 Etude de la restauration de la restauration « wild type » observée dans la souche $\Delta dus4 \Delta dus2$

Préalablement à la réalisation de ce mémoire, il a été observé au laboratoire que la délétion du gène *dus2* supprime le défaut de sporulation observé dans la souche  $\Delta dus4$ , (*dus2::kan<sup>R</sup> dus4::ble<sup>R</sup>*), aux niveaux phénotypique et moléculaire (expression de *spo6*, données RNAseq et Northern Blot non montrés). Cette constatation sous-entendait que le retrait d'une enzyme Dus empêchait l'expression du phénotype observé suite au retrait d'une autre enzyme Dus, fait interpellant qui tendait à prétendre une connexion entre ces deux enzymes, via la dihydrouridine.

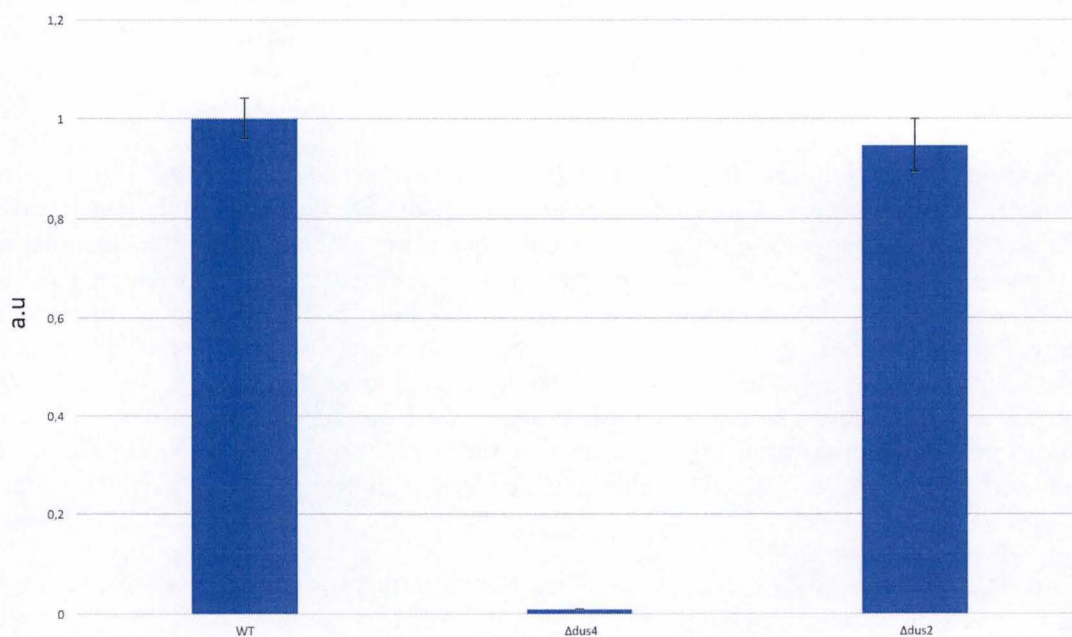
L'implication de l'activité catalytique de la protéine dihydrouridine synthase 2 dans cette restauration de phénotype a d'abord été investiguée. Une souche catalytiquement inactive pour la protéine Dus2, appelée *dus2<sub>DEAD</sub>*, a été créée et ce sur le même modèle de construction que la souche *dus4<sub>DEAD</sub>* (**Fig. 32**). Cette construction a ensuite été vérifiée par primer extension, avec en contrôle les souches wild type,  $\Delta dus2$ ,  $\Delta dus1-3-4$  (**Fig. 33**). Le fragment correspondant à la dihydrouridine associée à la dihydrouridine synthase 2, en position D19, se retrouve présent dans les souches wild type et  $\Delta dus1-3-4$  mais non dans les souches  $\Delta dus2$  et *dus2<sub>DEAD</sub>*. A noter la présence de bandes supplémentaires dans la souche  $\Delta dus2$ , correspondant peut-être à d'autres modifications post-transcriptionnelles qui arrêtent l'action de la reverse transcriptase. Cette donnée confirme le fait que l'activité catalytique de Dus2 est non-fonctionnelle dans le mutant *dus2<sub>DEAD</sub>*.

Une souche  $\Delta dus4 \text{ } dus2_{DEAD}$  a ensuite été créée par croisement à partir de la souche  $\Delta dus4 \Delta dus2$  et de la souche *dus2<sub>DEAD</sub>* (**Fig. 34**) afin d'ensuite vérifier par analyse phénotypique si la suppression de l'activité catalytique est suffisante pour restaurer le phénotype wild type.





**Figure 35. L'activité catalytique du la Dihydrouridine synthase 2 n'est pas nécessaire à la restauration du phénotype wild type.** Pourcentage de zygotes (bleu), asques (orange), asques présentant un défaut de sporulation (gris) 48h après l'induction de sporulation dans les souches wild type (WT),  $\Delta dus4$ , *dus2*DEAD et  $\Delta dus4dus2$ DEAD. Les barres d'erreur représentent l'écart-type pour trois répliques biologiques indépendants (représentant 600 cellules).



**Figure 36. qRT-PCR du gène *dus4* dans les souches wild type (WT),  $\Delta dus4$  et  $\Delta dus2$ .** Les barres d'erreur représentent la SEM pour trois répliques techniques, l'ARNm du *house keeping gene act1* a été utilisé pour la normalisation. a.u = arbitrary units (WT = 1).



Il s'avère que les souches *dus2*<sub>DEAD</sub> et  $\Delta$ *dus4 dus2*<sub>DEAD</sub> sporulent d'une façon similaire à la souche wild type, avec uniquement 14% d'asques déformés suite à l'induction de la sporulation (Fig. 35). Cela tend à appuyer le fait que l'activité catalytique de la protéine n'est pas nécessaire à la restauration du phénotype wild type observée dans la souche  $\Delta$ *dus4*  $\Delta$ *dus2*.

Ensuite, une hypothèse a été émise concernant une possible régulation compensatrice d'expression génique entre les gènes *dus2* et *dus4*. La diminution d'expression de *dus4* pourrait potentiellement être compensée par une augmentation de l'expression de *dus2*. Pour le vérifier, le niveau du transcrit de *dus4* a été quantifié dans une souche  $\Delta$ *dus2* (génotype : *dus2::kan<sup>R</sup>*). Par qRT-PCR (Fig. 36) et par Northern Bot (Fig. 26), il apparaît que l'expression de *dus4* n'est pas augmentée comparativement à une souche wild type dans la souche  $\Delta$ *dus2*. On peut alors exclure une régulation d'expression génique compensatrice entre les gènes *dus2* et *dus4*.

L'étude du défaut de sporulation observé dans la souche  $\Delta$ *dus4* de *S. pombe*, potentiellement corrélé à une perturbation de l'expression du gène *spo6*, et restauré à un phénotype wild type dans une souche  $\Delta$ *dus4*  $\Delta$ *dus2*, s'est arrêtée avec ces observations. Tous les points énumérés ci-dessus seront discutés dans la discussion, mais une conclusion peut déjà ici être faite concernant le rôle biologique potentiel de la dihydrouridine dans la sporulation chez *S. pombe*. Celle que le phénotype observé était finalement *Dus4*-indépendant, étant donné les résultats obtenus au fil des expériences réalisées.

#### IV.2 Préparation d'une analyse épitranscriptomique comparative de la dihydrouridine chez *Schizosaccharomyces pombe*, *Escherichia coli* et *Caenorhabditis elegans*

Afin d'ensuite permettre une étude fine de la disposition de la dihydrouridine au sein du transcriptome (comme le nécessite toute étude épitranscriptomique de modification post-transcriptionnelle), la seconde partie de ce mémoire se penche autour d'une technique de séquençage mise au point par Olivier Finet. En effet, dans le cadre de sa thèse sur l'étude épitranscriptomique de la dihydrouridine, Olivier Finet (doctorant en Sciences supervisant ce mémoire) a été amené à mettre au point une technique permettant de cartographier les positions de la dihydrouridine sur l'ensemble du transcriptome de *S. pombe*. Cette technique, appelée D-seq (pour dihydrouridine-sequencing), se base sur du séquençage d'ADNc obtenu après une rétro-transcription qui aura été bloquée ou non par la présence de dihydrouridine. Il s'agit du même principe qui est utilisé dans la technique de « primer extension » (voir introduction et première partie du mémoire). Pour rappel, après avoir extrait l'ARN, un traitement chimique de marquage de la dihydrouridine par de la rhodamine est réalisé. De cette manière, la rétro-transcription est bloquée quand la reverse transcriptase fait face au groupement chimique dihydrouridine-rhodamine. Des fragments d'ADNc de différentes tailles sont donc obtenus, en fonction de la présence ou non de dihydrouridine à la fin de ceux-ci. Et ce sont ces fragments d'ADNc de différentes tailles qui seront finalement séquencés. En fonction d'un ratio calculé pour chaque position du transcriptome, il sera alors possible de déterminer la présence ou non de dihydrouridine sur les gènes concernés par la position en question.

Après analyse fine des premiers résultats du D-seq réalisé sur un réplica d'ARN de *S. pombe* par Olivier Finet, 287 gènes codant pour des protéines sont ressortis comme étant dihydrouridylés, ainsi que 223 sites positionnés sur des ARN de transfert. Parmi ces gènes, cinq ARNm codant pour des sous-unités de l'octamère CCT portent au moins une dihydrouridine. Le CCT est une chaperonine formée par 8 sous-unités, et est active dans le cytosol des archéobactéries et des eucaryotes<sup>49</sup>.



Il assure le repliement correct de l'actine et la tubuline, protéines mères du cytosquelette, via différentes interactions en fonction des sous-unités concernées <sup>49</sup>. L'objectif de ce projet au laboratoire est de comprendre le rôle de la dihydrouridine dans les ARNm et d'établir sa prévalence au cours de l'évolution.

Pour y arriver, il a été décidé de réaliser du D-seq sur une espèce procaryote : *Escherichia coli* ainsi que sur une espèce eucaryote pluricellulaire : *Caenorhabditis elegans*, ces deux espèces étant déjà étudiées au laboratoire de GéMo (Génétique Moléculaire, UNamur). Pour rappel, des comparaisons phylogénétiques concernant les différentes dihydrouridine synthases <sup>16</sup> ou les différentes positions de dihydrouridine sur les ARN de transfert <sup>39, 15</sup> existent déjà. Dans ce cas-ci, ce serait la première fois qu'une comparaison phylogénétique concernant les différentes positions de dihydrouridine sur l'ensemble du transcriptome serait réalisée.

Le but de cette deuxième partie de mémoire était donc de préparer les étapes préliminaires à l'envoi d'échantillons d'*Escherichia coli* et de *Caenorhabditis elegans* au D-séquençage.

#### IV.2.1 Obtention, contrôles, culture et récolte des souches d'*Escherichia coli* et *Caenorhabditis elegans*

- *Escherichia coli*

Deux souches provenant de l'étude de Bishop *et al* <sup>12</sup> sur l'identification des dihydrouridine synthases des ARNt ont été utilisées. Une souche wild type (JC8679) et une souche dont les gènes codant pour les trois dihydrouridine synthases d'*E. coli* (*dusA*, *dusB*, *dusC*) ont été délétés (PS5107). Dans la souche PS5107, le gène codant pour DusA a été remplacé par une cassette de résistance au chloramphénicol, le gène codant pour DusB a été remplacé par une cassette de résistance à la kanamycine et le gène codant pour DusC a été remplacé par une cassette de résistance à la spectinomycine <sup>12</sup>. Ces trois résistances ont été vérifiées en faisant pousser la souche PS5107 ( $\Delta$ *dusABC*) sur un milieu solide comportant les antibiotiques en question. La souche poussant bien sur ces trois antibiotiques, il a été considéré que les gènes *dusA*, *dusB*, et *dusC* avait effectivement été délétés.

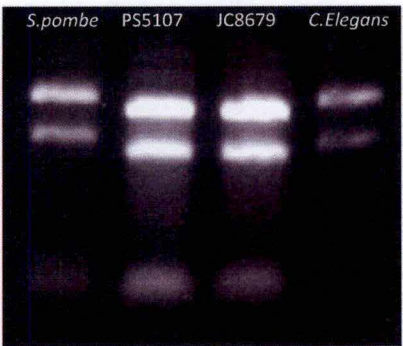
Les culots des deux souches JC8679 (WT) et PS5107 ( $\Delta$ *dusABC*) ont ensuite été récoltés après mise en culture liquide à 37°C, 100 mL de LB ont été utilisés pour un premier essai arbitraire. Dans l'étude de Bishop *et al* <sup>12</sup>, le gène *recBC* est surexprimé dans la souche WT (JC8679) à partir de laquelle a été créée la souche  $\Delta$ *dusABC* (PS5107) pour leur permettre une croissance efficace. Aucune anomalie de croissance n'a en effet été détectée lors de la mise en culture de ces souches (3h30 environ étaient nécessaire pour obtenir les bactéries à DO<sub>595nm</sub> 0,6-0,8 à partir d'une DO<sub>595nm</sub> de 0,25).

- *Caenorhabditis elegans*

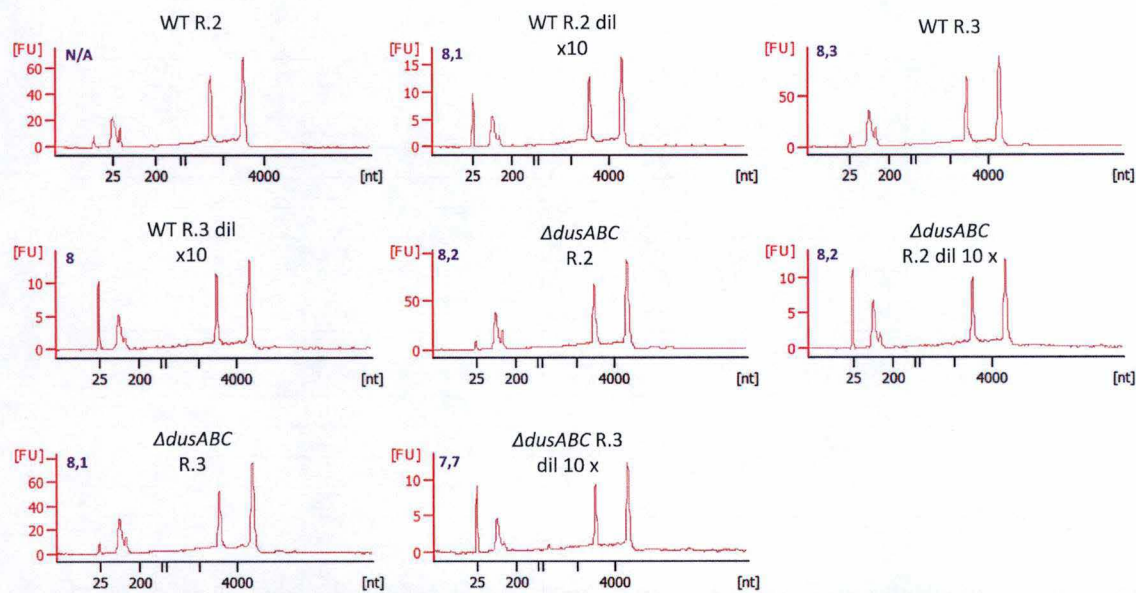
La souche wild type *Caenorhabditis elegans* N2 a été utilisée. Aucune souche délétée en enzymes Dus n'était disponible pour cette espèce (voir [www.wormbase.org](http://www.wormbase.org)). Etant donné la nécessité d'obtenir 30 µg d'ARN après extraction pour effectuer le rhodamine labelling (voir plus loin) et en se basant sur des données de récoltes précédentes de Fanélie Baueur (chercheuse en GéMo, UNamur), 200 000 vers minimum s'avéraient être potentiellement nécessaires pour chaque condition à séquencer. Pour chaque condition, 10 vers adultes portant en moyenne 200 embryons chacun ont d'abord été cultivés sur milieu solide. Après un cycle naturel complet d'environ 3 jours, les 10 x 200 adultes ont été « bleachés ».



Espèce	Souche-Génotype	Numéro réplica	Quantification (µg/µl)
<i>E.coli</i>	JC8679 - WT	1	3,686
<i>E.coli</i>	JC8679 - WT	2	3,485
<i>E.coli</i>	JC8679 - WT	3	3,726
<i>E.coli</i>	JC8679 - WT	4	3,845
<i>E.coli</i>	JC8679 - WT	5	3,974
<i>E.coli</i>	PS5107 - <i>dusABC</i>	1	3,673
<i>E.coli</i>	PS5107 - <i>dusABC</i>	2	3,880
<i>E.coli</i>	PS5107 - <i>dusABC</i>	3	3,814
<i>E.coli</i>	PS5107 - <i>dusABC</i>	4	3,776
<i>E.coli</i>	PS5107 - <i>dusABC</i>	5	3,905
<i>C.elegans</i>	N2 - WT	1	0,237
<i>C.elegans</i>	N2 - WT	2	0,332
<i>C.elegans</i>	N2 - WT	3	0,344
<i>C.elegans</i>	N2 - WT	4	0,534

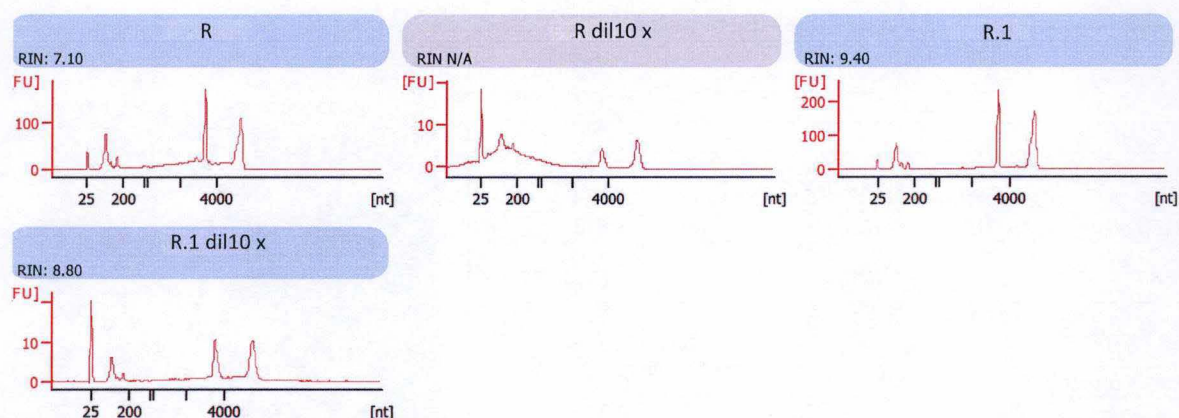


**Figure 37. Extraction des ARN des souches d'Escherichia coli et Caenorhabditis elegans.** Table : quantification de l'ARN total après extraction selon une méthode initialement destinée à l'extraction d'ARN de Schizosaccharomyces pombe. Au total 200 µL sont récoltés. Photographie : visualisation des ARN ribosomiques (haut) et des petits ARN (bas) des répliques n°1 (en gras dans la table) des extraits d'ARN récoltés. Une souche de Schizosaccharomyces pombe a été utilisée comme échantillon contrôle. ARNr de Schizosaccharomyces pombe : 25S (3400 nt) et 18S (1800 nt), ARNr de Escherichia coli : 23S et 16S, ARNr de Caenorhabditis elegans : 25S et 18S. Les répliques 2 et 3 d'Escherichia coli et le réplique 1 de Caenorhabditis elegans (violet) ont été analysés par bioanalyser (voir Fig. 38)



**Figure 38. Analyse Bioanalyser des ARN extraits d'Escherichia coli.** Les répliques 2 et 3 (voir figure 37, violet) ont été analysés, une dilution 10x a été réalisée pour chaque réplica. 1 µL d'ARN à concentration 3,5 ng/µL et 0,35ng/µL respectivement a été chargé sur la puce avant analyse. Le pic situé à 25 nt correspond à une calibration. Le pic situé à environ 100 nt correspond aux petits ARN. Les pics situés à 1000 et à 4000 nt correspondent aux ARN ribosomiques. Entre les petits ARN et les ARNr se situent les ARN messagers. Si le signal reste plat, alors l'échantillon est de bonne qualité, c'est le cas ici. De plus, les RIN (RNA integrated number) écrits en violet pour chaque échantillon et correspondant à la proportion d'ARNr quantifiée par rapport à l'ARN total se situent dans un range acceptable (idéal : 10). L'analyse de la puce a été réalisée par une quantification propre aux procaryotes. R= réplica.





**Figure 39. Analyse Bioanalyzer des ARN extraits de *Caenorhabditis elegans*.** Le réplica 1 (voir figure 37, violet, R.1) et un autre réplica test non répertorié dans ce mémoire ont été analysés, une dilution 10x a été réalisée pour chaque réplica. 1 µL d'ARN à concentration 3,5 ng/µL et 0,35ng/µL respectivement a été chargé sur la puce avant analyse. Le pic situé à 25 nt correspond à une calibration. Le pic situé à environ 100 nt correspond aux petits ARN. Les pics situés à 3500 et à 5000 nt correspondent aux ARN ribosomiques. Entre les petits ARN et les ARNr se situent les ARN messagers. Si le signal reste plat, alors l'échantillon est de bonne qualité, c'est le cas ici. De plus, les RIN (RNA integrated number) correspondant à la proportion d'ARNr quantifiée par rapport à l'ARN total se situent dans un range acceptable (idéal : 10). L'analyse de la puce a été réalisée par une quantification propre aux eucaryotes. R= réplica.

Espèce	Souche-Génotype	Labelling	Numéro réplica	Quantification dilution 10x (µg/µl)
<i>E.coli</i>	JC8679 - WT	R +	1	0,406
<i>E.coli</i>	JC8679 - WT	R -	1	0,474
<i>E.coli</i>	JC8679 - WT	R +	2	0,544
<i>E.coli</i>	JC8679 - WT	R -	2	0,298
<i>E.coli</i>	PS5107 - <i>dusABC</i>	R +	1	0,604
<i>E.coli</i>	PS5107 - <i>dusABC</i>	R -	1	0,508
<i>E.coli</i>	PS5107 - <i>dusABC</i>	R +	2	0,367
<i>E.coli</i>	PS5107 - <i>dusABC</i>	R -	2	0,354
<i>C.elegans</i>	N2 - WT	R +	1	0,214
<i>C.elegans</i>	N2 - WT	R -	1	0,711
<i>C.elegans</i>	N2 - WT	R +	2	0,446
<i>C.elegans</i>	N2 - WT	R -	2	0,412
<i>C.elegans</i>	N2 - WT	R +	3	0,285
<i>C.elegans</i>	N2 - WT	R -	3	0,274

**Figure 40. Quantification de l'ARN total après rhodamine labelling.** 2 réplicas pour chaque souche d'*Escherichia coli* sont ici représentés et 3 réplicas pour *Caenorhabditis elegans*. Les ARN de départ ont été extraits selon un protocole initialement destiné à l'extraction d'ARN de *Schizosaccharomyces pombe*. Au total 7 µL sont récoltés. La quantification a été effectuée à partir d'une dilution 10x de chaque échantillon dans un volume final de 10 µL. Chaque réplica d'ARN extrait a subi un traitement d'incorporation de la rhodamine (R +) et un traitement où la rhodamine n'est pas incorporée (R -).



Le bleaching est une pratique qui consiste à dissoudre les vers adultes gravides dans une solution à base d'eau de Javel, afin de ne récupérer que les embryons, pour ensuite pouvoir synchroniser l'état de croissance des vers lors du cycle suivant. Environ 10 embryons sont récoltés par adultes après bleaching. Les 20 000 (10 x 200 x 10) embryons ont ensuite été cultivés et nourris en milieu liquide et un second bleaching a été réalisé une fois les embryons devenus adultes. Les 200 000 (20 000 x 10) embryons ont ensuite poussés jusqu'au stade L4 avant d'être récoltés et culotés.

#### IV.2.2 Tests d'extraction d'ARN et analyses réalisés sur les souches d'*Escherichia coli* et *Caenorhabditis elegans* récoltées

Après récolte des culots d'*Escherichia coli* et *Caenorhabditis elegans*, il a fallu extraire leur ARN afin d'ensuite pouvoir effectuer le marquage de la dihydrouridine par la rhodamine.

Afin de pouvoir comparer les résultats du D-seq à réaliser sur *E. coli* et *C. elegans* avec les résultats déjà obtenus par Olivier Finet sur *S. pombe*, il était préférable d'effectuer les mêmes manipulations. C'est pourquoi le protocole d'extraction d'ARN initialement destiné à *S. pombe* a été testé (voir matériels et méthodes). *E. coli* possédant un peptidoglycan, *C. elegans* une cuticule et *S. pombe* possédant une paroi qui s'avère plus épaisse, il a été supposé que l'extraction devait autant être réalisable sur *E. coli* et *C. elegans*. Les quantifications obtenues sont référencées dans la figure 37 (**Fig. 37**). Avec cette méthode, il apparaît finalement aisé de récolter en suffisance de l'ARN chez *E. coli*, mais beaucoup moins d'ARN est récoltable chez *C. elegans*. Etant donné que la récolte finale d'ARN se fait dans des échantillons de 200 µL et que 30 µg d'ARN sont nécessaires pour démarrer le marquage par la rhodamine (rhodamine labelling), les quantités initiales d'*E. coli* et de *C. elegans* utilisées (rappel : 100 mL de LB d'*E. coli* et 200 000 vers de *C. elegans*) s'avèrent cependant être suffisantes.

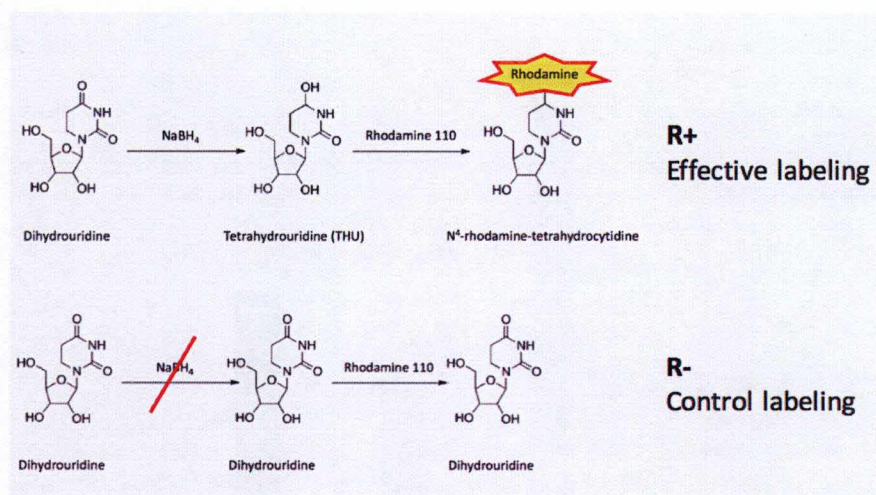
La qualité des ARN extraits a ensuite été vérifiée. Premièrement, les extraits totaux d'ARN ont été séparés par gel d'agarose et observés par bromure d'éthidium. De la sorte, les ARNs les plus abondants sont clairement identifiables et donnent un aperçu de la qualité de l'extrait (**Fig. 37**). Deuxièmement, le profil de ces ARN d'*E. coli* et de *C. elegans* a été analysé par une quantification microfluidique de l'ARN via la technique de Bioanalyzer (**Fig. 38** et **Fig. 39**).

Etant donné la qualité des ARN qui s'avèrent non dégradés (voir **Fig. 37**) et en proportion suffisante (RIN, RNA integrated number, élevé, voir **Fig. 38** et **Fig. 39**), la méthode d'extraction d'ARN utilisée a été jugée satisfaisante.

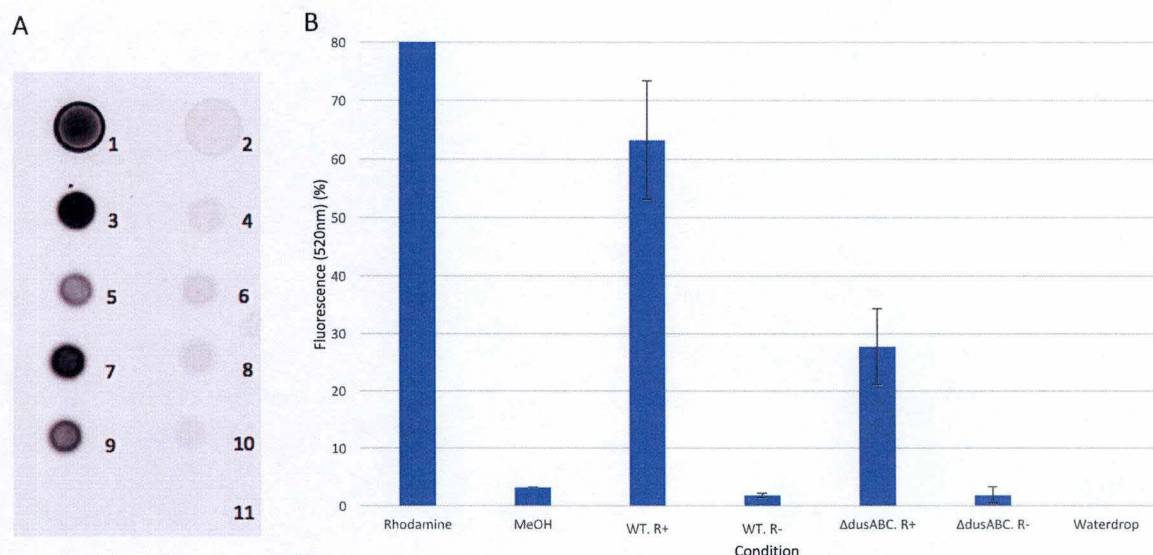
#### IV.2.3 « Rhodamine labelling » sur les ARN extraits d'*Escherichia coli* et *Caenorhabditis elegans*

Afin de pouvoir distinguer au séquençage la position de la dihydrouridine sur les ADNc séquencés à partir d'ARN, un traitement chimique à base de rhodamine a été réalisé. L'ARN rhodaminisé a ensuite été quantifié. Les quantifications des essais finaux sont répertoriées dans la figure 40 (**Fig. 40**). La variation de quantité d'ARN par rapport à la quantification avant marquage par la rhodamine est jugée acceptable, étant donné les résultats similaires obtenus après de nombreux essais de rhodamine labelling effectués par Olivier Finet au cours de sa thèse. Si la quantification d'ARN rhodaminisé s'était avérée beaucoup plus supérieure à la quantification d'ARN de départ, cela aurait pu supposer une probable contamination des échantillons.



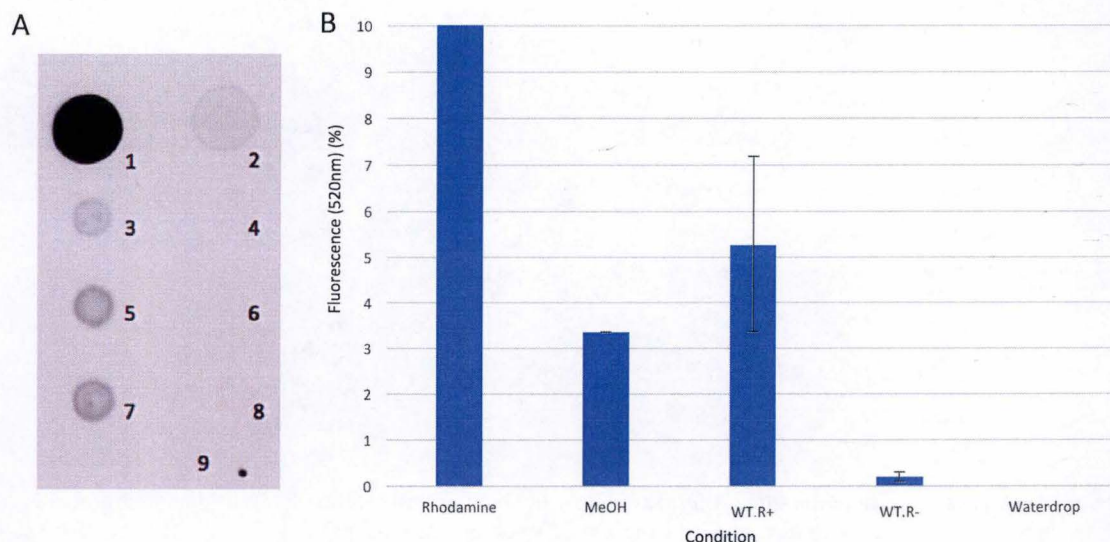


**Figure 41. Détails du marquage de la dihydrouridine par la rhodamine.** La dihydrouridine doit d'abord être réduite en tétrahydrouridine afin de pouvoir incorporer efficacement la rhodamine. Ici, du  $\text{NaBH}_4$  est utilisé pour la réduction (échantillons R +). Dans les échantillons contrôles (R-), la dihydrouridine n'est pas réduite et la rhodamine ajoutée ne peut donc pas se fixer sur la dihydrouridine. Figure d'Olivier Finet.



**Figure 42. Dot Blot assay des échantillons d'ARN d'*Escherichia coli* après rhodamine labelling.** Des gouttes de  $3\mu\text{L}$  à concentration  $0,5\mu\text{g}/\mu\text{L}$  ont été placées sur une membrane de nylon et la fluorescence visualisée et quantifiée avec le logiciel Imagquant. **A.** Photographie des gouttes. 1 : rhodamine 2 : Méthanol 3 : JC8679 – WT réplique 1 R+ 4 : JC8679 – WT réplique 1 R- 5 : JC8679 – WT réplique 2 R+ 6 : JC8679 – WT réplique 2 R- 7 : PS5107 – *dusABC* réplique 1 R+ 8 : PS5107 – *dusABC* réplique 1 R- 9 : PS5107 – *dusABC* réplique 2 R+ 10 : PS5107 – *dusABC* réplique 2 R- 11 : eau. **B.** Pourcentage de fluorescence émise par la rhodamine incorporée en fonction de la condition, moyenne des deux répliques.





**Figure 43. Dot Blot assay des échantillons d'ARN de *Caenorhabditis elegans* après rhodamine labelling.** Des gouttes de 3 $\mu$ L à concentration 0,5  $\mu$ g/ $\mu$ L ont été placées sur une membrane de nylon et la fluorescence visualisée et quantifiée avec le logiciel Imagquant. **A.** Photographie des gouttes. 1 : rhodamine 2 : Méthanol 3 : N2 – WT réplique 1 R+ 4 : N2 – WT réplique 1 R- 5 : N2 – WT réplique 2 R+ 6 : N2 – WT réplique 2 R- 7 : N2 – WT réplique 3 R+ 8 : N2 – WT réplique 3 R- 9 : eau. **B.** Pourcentage de fluorescence émise par la rhodamine incorporée en fonction de la condition, moyenne des trois répliques.



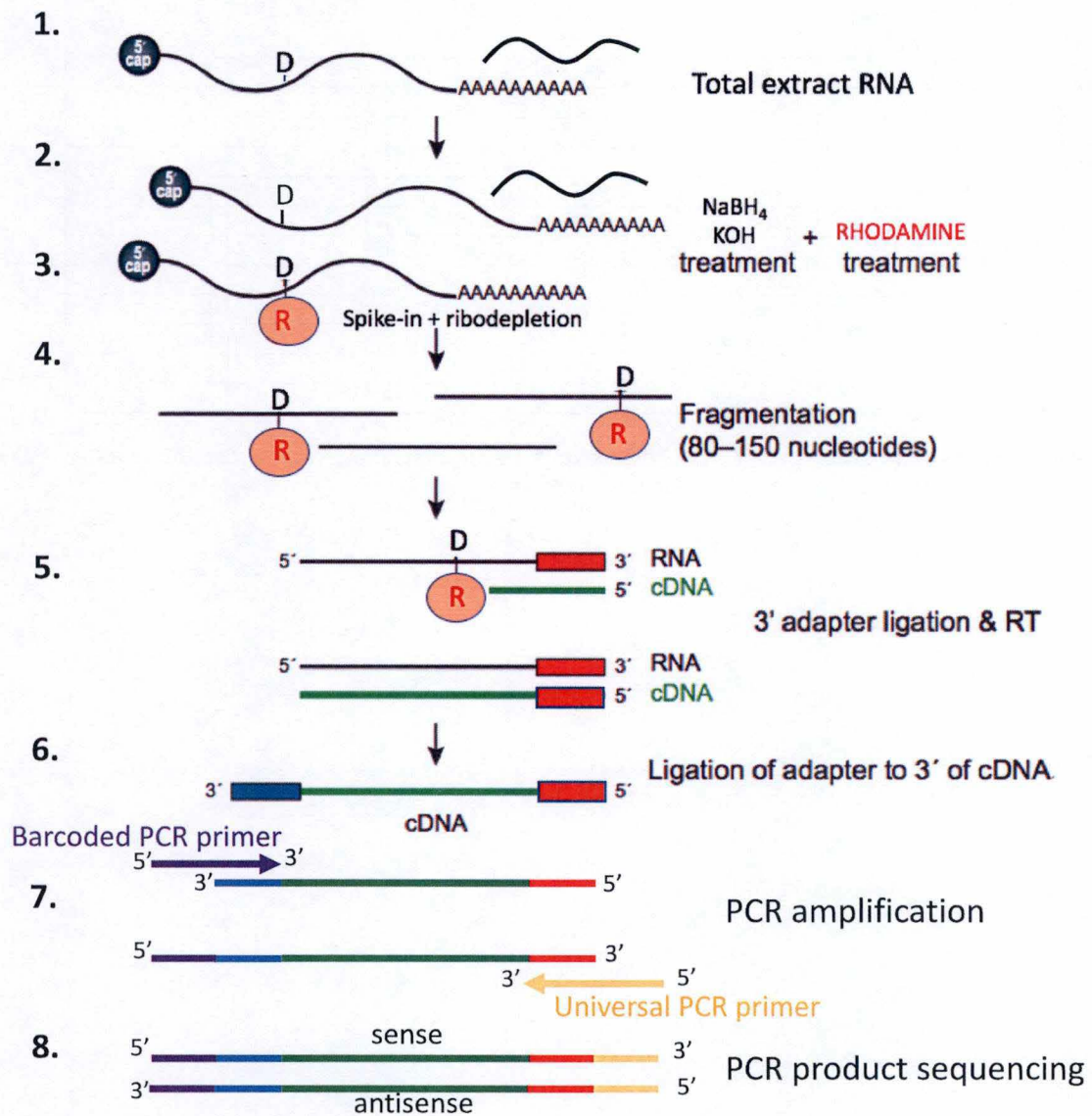
Le marquage de la dihydrouridine par la rhodamine a été vérifié par une technique de « dot blot assay ». Cette méthode est basée sur la mesure de la fluorescence émise par des gouttes des échantillons obtenus après le rhodamine labelling. La rhodamine étant un fluorophore, il est possible de quantifier la fluorescence émise par les échantillons marqués. Chaque marquage de souche d'*E. coli* et de *C. elegans* (échantillons « R+ ») a été contrôlé par une souche de même génotype (WT ou *ΔdusABC*) mais non marquée par la rhodamine (échantillons « R- »). De manière détaillée, la dihydrouridine doit d'abord être réduite pour pouvoir incorporer la rhodamine (Fig. 41). Ici, un traitement à base d' $\text{NaBH}_4$  dilué dans du KOH est réalisé. Dans les échantillons R-, uniquement du KOH est utilisé de sorte que la dihydrouridine ne soit pas réduite et donc incapable d'être liée par la rhodamine une fois cette dernière ajoutée dans l'échantillon.

Concernant les souches d'*E. coli*, il est possible de distinguer une diminution de fluorescence dans la souche dépourvue de dihydrouridine synthase (PS5107), en comparaison avec la souche wild type (JC8679), ce qui confirme le marquage de la dihydrouridine (Fig. 42). De la rhodamine reste cependant quantifiable, ce qui est attendu puisque d'autres modifications de l'ARN sont susceptibles d'être marquées<sup>36</sup>. Concernant *C. elegans*, il est possible de vérifier qu'il y a bien eu plus d'incorporation de la rhodamine dans les échantillons R+, en comparaison avec les échantillons R- (Fig. 43), ce qui valide la méthode de rhodamine labelling chez *C. elegans* également. Toutefois, l'absence de mutants disponibles pour les enzymes Dus ne permet pas d'estimer la contribution des dihydrouridines dans le marquage.

#### IV.2.4 Etapes du « D-seq »

Une fois les échantillons d'ARN rhodaminisés ou non obtenus en quantité suffisante (besoin de 2 µg), la première étape à réaliser est une étape de ribodéplétion. Cela consiste à se débarrasser des ARN ribosomiques qui sont des ARN présents en très grande quantité dans chaque extrait total d'ARN. Etant donné la quantité importante de dihydrouridine présente sur les ARN ribosomiques<sup>50</sup>, cela affine les résultats de séquençage sur ARNm que de s'en débarrasser. Pour cela, un kit de ribodéplétion utilisable sur des échantillons de *C. elegans* (eucaryote pluricellulaire) a dû être commandé. Ce kit commandé s'avérerait être aussi utilisable pour la ribodéplétion des ARN d'*E. coli* (procaryote). Les échantillons d'ARN rhodaminisés ribodéplétés sont ensuite concentrés et la ribodéplétion est à vérifier par technique Bioanalyzer. Dans ce cas, on s'attend à perdre les pics situés à 1000 nt et 4000 nt correspondant aux ARNr. Ensuite, une fragmentation des échantillons est nécessaire pour permettre une bonne rétro-transcription de l'entièreté du transcriptome. Un traitement DNase est ensuite réalisé après élution afin d'éviter toute dégradation de l'ADNc qui sera produit après rétro-transcription. Par la suite, une première ligation est réalisée avec un adaptateur ARN placé en extrémité 3' pour ensuite permettre de réaliser la rétro-transcription avec le primer complémentaire de l'adaptateur utilisé (dans le sens 5'-3'). Le fait d'ajouter des étapes de ligation d'oligonucléotides (d'abord sur l'ARN pour la rétro-transcription, ensuite sur l'ADNc pour la synthèse du second brin) diffère des protocoles usuels de RNAseq. En effet, dans le cas classique, la rétro-transcription engendre, à partir du même fragment, des ADNc de tailles diverses car des « random primers » sont utilisés. C'est à partir de ces fragments – qui peuvent être considérés comme aléatoirement générés – que le second brin d'ADN est synthétisé. Or si celui-ci est également produit de manière aléatoire (comme dans le RNAseq classique), cela signifie qu'à partir d'un même fragment d'ARN de départ, l'ADN obtenu ne peut pas être caractérisé précisément quant à son extrémité 3'. C'est cependant cette caractéristique qui est ici utilisée, dans le D-seq, pour déterminer une position comme une dihydrouridine ou pas.





**Figure 44. Schématisation des étapes du D-sequencing des échantillons d'ARN d'*Escherichia coli* et *Caenorhabditis elegans*.** 1. L'ARN total est extrait selon un protocole d'extraction ARN initialement destiné à *Schizosaccharomyces pombe*. 2. Rhodamine labelling 3. Ajout de spike-in et ribodéplétion. 4. Fragmentation des ARN rhodaminisés et ribodéplétés 5. Ajout d'un adaptateur ARN en 3' et reverse-transcription. Les cDNA obtenus (vert) auront une taille différente en fonction de la présence ou non de dihydrouridine marquée par la rhodamine lors du rhodamine labelling. 6. Ajout d'un adaptateur ADN en extrémité 3'. 7. Amplification PCR à l'aide d'un barcoded PCR primer (violet) propre à chaque librairie et d'un universal PCR primer (jaune). 8. Produit PCR à séquencer par la méthode Illumina sequencing. Figure adaptée de Schwartz *et al.* <sup>35</sup>



<u>Numéro librairie</u>	<u>Espèce</u>	<u>Souche-Génotype</u>	<u>Réplicas</u>	<u>Labelling</u>
1	<i>E.coli</i>	JC8679 - WT	1	R +
2	<i>E.coli</i>	JC8679 - WT	1	R -
3	<i>E.coli</i>	JC8679 - WT	2	R +
4	<i>E.coli</i>	JC8679 - WT	2	R -
5	<i>E.coli</i>	PS5107 - <i>dusABC</i>	1	R +
6	<i>E.coli</i>	PS5107 - <i>dusABC</i>	1	R -
7	<i>E.coli</i>	PS5107 - <i>dusABC</i>	2	R +
8	<i>E.coli</i>	PS5107 - <i>dusABC</i>	2	R -
9	<i>C.elegans</i>	N2 - WT	1	R +
10	<i>C.elegans</i>	N2 - WT	1	R -
11	<i>C.elegans</i>	N2 - WT	2	R +
12	<i>C.elegans</i>	N2 - WT	2	R -

**Figure 45. Tableau représentatif des 12 librairies à D-séquencer.** 2 réplicas seront séquencés pour chaque souche, chaque réplica étant marqué ou non par la rhodamine.



Autrement dit, en considérant un fragment d'ARN  $x$ , la ligation d'un adaptateur en son extrémité 3' permet d'engendrer des ADNc qui, pour le dire simplement, partent tous du même endroit (mais ne s'arrêteront pas tous au même endroit, en fonction de positions où l'enzyme sera bloquée ou arrêtée). La seconde ligation (sur l'ADNc) permet, quant à elle, de s'assurer que l'extrémité 3' de l'ADNc (correspondant au site précis de blocage ou d'arrêt de la rétro-transcription sur l'ARN) est bien prise en compte dans la production d'un ADN double brin qui est ce qui sera véritablement séquencé. Les cDNA de différentes tailles (en fonction de la présence de dihydrouridine ou non) obtenus sont ensuite nettoyés et le reste d'ARN dégradé. Suite à la seconde ligation, c'est un barcoded primer qui est utilisé pour permettre l'amplification PCR des cDNA obtenus (explications voir point IV.2.5). Les différentes étapes ainsi que le produit final envoyé au séquençage sont schématisés dans la figure 44 (Fig. 44).

#### IV.2.5 Préparation des librairies à envoyer au séquençage

Plusieurs contrôles sont nécessaires afin de vérifier la bonne fonctionnalité de la méthode. Un premier contrôle des résultats est l'envoi de deux répliques biologiques pour la même condition (*E.coli* WT ou *E.coli*  $\Delta$ *dusABC* ou *C.elegans* WT). Un second contrôle est l'envoi d'échantillons à la fois rhodaminisés et non rhodaminisés. Cela revient à 2 échantillons (un rhodaminisé et un non rhodaminisé) pour chaque réplique. Vu que l'on a deux échantillons (R+ et R-), deux répliques pour chaque condition et trois conditions différentes en tout (*E.coli* WT, *E.coli*  $\Delta$ *dusABC*, *C.elegans* WT), 12 librairies sont donc à préparer au total (résumé voir Fig. 45).

Pour ensuite contrôler la méthode de quantification de la dihydrouridine, des spike-in ARN sont à ajouter dans chaque échantillon. Il s'agit d'une séquence d'ARN connue contenant une dihydrouridine ou une uridine à une même position particulière. Cette séquence d'ARN est obtenue à partir d'une séquence d'ADN (voir Schwartz *et al* <sup>15</sup>) amplifiée par PCR puis transcrite in vitro avec l'ajout de dNTP (rGTP, rCTP, rATP, rUTP ou rDTP). Deux échantillons d'ARN spike-in, l'un contenant de la dihydrouridine et l'autre de l'uridine sont ainsi produits et utilisés dans chaque librairie. Dans chaque échantillon à D-séquencer, un certain pourcentage de spike-in dihydrouridylé et un certain pourcentage de spike-in uridylé (spike-in D/ spike-in U) sont ensuite ajoutés, en plus de l'ARN à séquencer. 12 pourcentages différents sont envisageables vu que 12 librairies différentes sont à séquencer. Après séquençage, reconnaissance de la séquence connue de spike-in, et calcul du ratio, il est possible de vérifier que l'on quantifie un pourcentage équivalent de spike-in D/ spike-in U séquencé par rapport au pourcentage de spike-in D/ spike-in U ajouté dans chaque librairie.

Afin de différencier les librairies entre elles lors de l'analyse des résultats des séquençages, des primer « PCR barecoded » différents sont utilisés pour chaque librairie. Ces primers sont utilisés lors de l'amplification PCR des cDNA finaux obtenus en fin de préparation de librairie. Ils sont hybridés à l'extrémité des séquences de cDNA à séquencer (Fig. 44). Dans ce cas-ci, 12 barcoded PCR primer sont nécessaires.

Les librairies envoyées sont lues sur des flowcell. Il s'agit de plateformes de séquençage à haut-débit pourvues d'adaptateurs ADN sur lesquels s'hybrident les cDNA sens et anti-sens à séquencer. Ici, la méthode de séquençage par synthèse « Illumina sequencing thecnology » est utilisée pour séquencer les librairies envoyées.



Cette méthode est basée sur une amplification des séquences hybridées sur la flowcell et leur lecture au moyen de traceurs fluorescents (méthode de Sanger <sup>51</sup>). Chaque flowcell dispose d'une certaine capacité de lecture en termes de paires de base : 300 millions de reads sont envisageables par flowcell. En fonction de la taille du génome de l'espèce envoyée et du nombre de librairies disposé sur chaque flowcell, chaque génome pourra être lu un certain nombre de fois, on parle de « coverage ». Etant donné que les génomes de trois espèces différentes sont ici à séquencer et que 12 librairies sont à envoyées, les coverages seront donc différents pour chaque espèce. Ce point est à prendre en compte lors du design des flowcell (par rapport au nombre de librairies qui y sera disposé) et de l'analyse des résultats. Pour information, *Schizosaccharomyces pombe* possède un génome de 12 millions de paires de base, *Escherichia coli* 5 millions de paires de base et *Caenorhabditis elegans* 100 millions de paires de base. Il a été décidé qu'une flowcell sera utilisée pour les 8 librairies d'*E.coli* et une autre pour les 4 librairies de *C.elegans*. De cette façon, le coverage du génome d'*E.coli* sera d'environ 7,5 ( $300 : (8 \times 5)$ ) et de 0,75 pour le génome de *C.elegans* ( $300 : (4 \times 100)$ ).

Au terme de ce mémoire, après avoir effectué les tests préliminaires à l'envoi des échantillons au D-seq et après avoir envisagé la préparation des librairies, la préparation du D-seq s'est arrêtée à la récolte des échantillons finaux d'*Escherichia coli* et *Caenorhabditis elegans*. Olivier Finet effectuera le D-sequencing des ARN finaux extraits d'*Escherichia coli* et de *Caenorhabditis elegans* et comparera les résultats obtenus lors de la suite de sa thèse sur la dihydrouridine.



## V. Discussion



## V. Discussion

Le but de ce mémoire était d'étudier une modification post-transcriptionnelle encore peu décrite : la dihydrouridine. Pour y parvenir, le mémoire s'est articulé en deux parties : une relative à l'étude du rôle biologique de la dihydrouridine et l'autre concernant la mise au point d'une technique de cartographie de cette modification à l'échelle du transcriptome.

### V.1 Implication de la dihydrouridine synthase Dus4 dans la sporulation chez *Schizosaccharomyces pombe*

La première partie des travaux de ce mémoire s'est basée sur un constat préalablement fait au laboratoire : un mutant de délétion de *S. pombe* pour le gène codant pour la dihydrouridine synthase 4 qui présentait un défaut de sporulation. Plutôt que de générer 4 gamètes (spores) uniformes par méiose, le mutant produit des spores dont le nombre et la morphologie sont altérés. De plus, l'analyse transcriptomique de cette souche, a révélé une sous-expression du gène de sporulation *spo6*. De manière très inattendue, la délétion additionnelle de *dus2* supprime ce défaut, suggérant une interaction génétique intéressante à étudier.

La première étape a été d'approfondir l'étude du défaut de sporulation et de la diminution d'expression du gène *spo6* dans la souche  $\Delta dus4$  par analyse phénotypique, qRT-PCR et Northern Blot. Les résultats confirment clairement le constat initial, notamment avec une diminution d'expression de *spo6* plus marquée 10 heures après l'induction de la sporulation.

Une seconde construction *dus4::ble<sup>R</sup>* a été réalisée (souche  $\Delta dus4_{BIS}$ ) afin de confirmer le défaut de sporulation. Le défaut est moins important dans cette souche comparativement à la première construction, mais il reste néanmoins clair (44% de « sporulation defective cells »). La différence pourrait être expliquée par un polymorphisme entre les souches ou par une variation aléatoire, lors des comptages, aucune analyse statistique n'ayant été réalisée dans le à ce stade du travail.

Afin de comprendre en profondeur la cause du défaut de sporulation et avant de le relier directement à la perte du gène *dus4* ou de la protéine Dus4, plusieurs souches ont été construites. En effet, l'environnement génomique de *dus4* est dense et l'introduction d'une cassette de résistance à un antibiotique n'est sans doute pas anodine pour la régulation de l'expression de ces gènes voisins.

Une première souche a été construite dans le but d'inactiver l'enzyme Dus4 (souche *dus4<sub>DEAD</sub>*) avec un minimum d'effet sur la protéine elle-même. Cette construction a été vérifiée par la réalisation d'un primer extension. Dans ce primer extension, l'absence d'un arrêt d'extension dépendant de Dus4 dans la souche *dus4<sub>DEAD</sub>* confirme l'absence d'activité catalytique de la protéine Dus4 dans cette construction. Cette expérience montre aussi que l'arrêt est beaucoup plus marqué dans la souche contrôle  $\Delta dus1-2-3$  que dans la wild type. Une explication potentielle est le fait que la conformation de l'ARN de transfert dans la souche  $\Delta dus1-2-3$  serait beaucoup plus lâche que dans une souche WT (étant donné l'absence de dihydrouridine pour aider à la conformation propre repliée de l'ARNt), ce qui faciliterait la réaction d'extension. De ce fait, la reverse transcriptase aurait moins de possibilités d'arrêt intempestifs avant de rencontrer la dihydrouridine synthétisée par Dus4 dans la souche  $\Delta dus1-2-3$  et engendrerait plus de fragments d'ARN de taille de 41nt, donc une bande plus apparente au primer extension.



L'analyse par comptages cellulaires de *dus4*<sup>DEAD</sup> montre que 20 % de sporulation defective cells sont présents. Par qRT-PCR ou Northern Blot, aucune diminution d'expression du gène *spo6* n'est apparente dans cette souche. L'écart-type des comptages pris en compte, il peut être supposé que ce n'est pas la perte de l'activité catalytique de la protéine Dus 4 dans la souche *Δdus4* qui soit responsable du défaut de sporulation observé. Le défaut de sporulation quand même observé dans les 20 % d'asques défectueux pourrait être dû au fait que d'autres ARN messagers impliqués dans le cycle méiotique (autres que *spo6*) soient dihydrouridylés par Dus4. En effet, lors du D-seq réalisé précédemment chez *S. pombe* par Olivier Finet, de nombreux autres ARN messagers étaient apparus comme étant dihydrouridylés. La perte d'activité catalytique de la protéine Dus4 pourrait donc provoquer un effet modéré sur la sporulation, mais d'une amplitude nettement plus faible que dans le mutant de délétion.

Pour déterminer si la présence de la protéine, indépendamment de son activité, était importante pour la sporulation, une souche dans laquelle 8 codons STOP ont été introduits dans la phase ouverte de lecture pour terminer précocement la traduction de l'ARNm *dus4* a été construite. La construction a d'abord été vérifiée par primer extension : l'absence de la bande correspondant à l'activité catalytique de la protéine Dus4 a bien été vérifiée. La réalisation d'un Western Blot avec des souches TAP-taguées confirme aussi l'absence de protéine Dus4 dans la souche *dus4*<sub>STOP</sub>-TAP, bien que le contrôle de charge ne soit pas judicieux puisque les conditions de croissance en milieu riche rendent la protéine Spo6-TAP difficilement détectable. Une analyse phénotypique a montré que la souche *dus4*<sub>STOP</sub> sporule de façon similaire à une souche WT. Ce n'est donc pas la perte de la protéine Dus 4 (dans son entièreté ou uniquement catalytiquement) dans la souche *Δdus4* qui provoque le défaut de sporulation observé.

Une autre hypothèse a ensuite été émise : celle de l'importance de la perte d'un transcrit émanant du locus *dus4* dans la souche *Δdus4* dans le défaut de sporulation. Pour la vérifier, la présence du transcrit dans la souche *dus4*<sub>STOP</sub> a été testée. En effet, vu que la souche *dus4*<sub>STOP</sub> sporule d'une manière wild type, si le transcrit n'y est plus présent, cela sous-entend qu'il n'est pas nécessaire à une bonne sporulation. Et inversement, si l'ARNm *dus4* est toujours bien présent dans la souche *dus4*<sub>STOP</sub>, c'est potentiellement sa perte dans la souche *Δdus4* qui peut être responsable du défaut de sporulation. Etant donné qu'il existe dans la cellule un mécanisme de contrôle qui élimine les transcrits contenant des codons de terminaison prématurés (NMD ; nonsense-mediated decay)<sup>52</sup>, l'ARNm *dus4*<sub>STOP</sub>, contenant 8 codons stop, pouvait être ciblé par cette voie et dégradé. Le but de ce mécanisme au sein de la cellule est d'éviter la formation de protéines potentiellement aberrantes ou toxiques pour la cellule<sup>52</sup>. Par quantification qRT-PCR et par Northern Blot, l'ARNm *dus4* apparaît bien présent dans la souche *dus4*<sub>STOP</sub>. Par qRT-PCR, il est possible de mettre en évidence une différence d'expression comparativement à la souche WT (0,56 versus 1), mais cela n'est pas visualisable par Northern Blot. Cette différence peut être expliquée par l'action de la NMD. A noter que le Northern Blot *dus4* n'étant pas brin spécifique, un deuxième ARN est marqué. Il s'agit ici de l'ARN situé sur le brin complémentaire du gène *dus4* : le long ARN non codant *SPNCRNA.1212.1*. Sur le site [www.pombase.org](http://www.pombase.org), *SPNCRNA.1212.1* est référencé avec une taille de 2905 nt. Cependant, ici, la taille de l'ARN marqué est d'environ 1700 nt.

La position précise de la sonde radioactive et le fait que les ARN non codants ne sont pas toujours correctement annotés dans la littérature du fait de l'absence de signaux de traduction clairs (ATG et STOP), il est probable que plusieurs ARNs non codant correspondent à *SPNCRNA.1212.1*. Nous concluons donc qu'au moins un ARN non codant d'une taille estimée à 1700 nucléotides est exprimé en anti-sens à *dus4*.



Une première hypothèse peut donc ici être émise concernant le défaut de sporulation observé dans la souche  $\Delta dus4$  : celle de l'impact de l'absence du transcrit *dus4*. Pour la vérifier, il serait judicieux de quantifier l'expression de *spo6* dans la souche *dus4*<sub>STOP</sub>. Cependant, rien n'est connu sur le mécanisme potentiel par lequel la perte de ce transcrit pourrait impacter la gamétogenèse chez la levure.

Introduire une cassette de résistance (gène contenant un promoteur propre) dans un génome peut être perturbant pour l'environnement génomique, notamment en perturbant la transcription des gènes voisins. L'impact de l'introduction de la cassette de résistance *ble*<sup>R</sup> dans la construction de la souche  $\Delta dus4$  a ensuite été investigué. Premièrement, une souche contenant un autre gène remplacé par la même cassette de résistance (souche *hst2::ble*<sup>R</sup>) a été mise en conditions de sporulation et des comptages cellulaires ont été faits. Etant donné que la souche *hst2::ble*<sup>R</sup> possède un phénotype comparable à celui d'une souche wild type, ce n'est pas la cassette de résistance *ble*<sup>R</sup> seule qui est responsable du défaut de sporulation. Ensuite, un autre marqueur (le gène *ura4*) a été utilisé dans la construction *dus4::ura4* afin de vérifier l'impact de la délétion du gène *dus4* autrement que par la cassette *ble*<sup>R</sup>. Cependant, ici, toute la CDS de *dus4* n'est pas remplacée par le gène *ura4*, ce qui implique la présence de deux promoteurs : celui de *dus4* et celui de *ura4*. Etant donné l'absence de défaut de sporulation dans la souche *dus4::ura4*, il est possible que cette construction conserve un niveau d'expression normal de l'ARN non codant antisense. Quoi qu'il en soit, cette construction confirme que la protéine Dus4 n'est pas requise pour la sporulation. Finalement, une troisième souche contenant un autre marqueur de résistance à un antibiotique qui remplace *dus4* a été construite (*dus4::hph*<sup>R</sup>). Dans cette souche, toute la CDS de *dus4* est remplacée, tout comme dans la souche initiale  $\Delta dus4$  (*dus4::ble*<sup>R</sup>). Le défaut de sporulation reste apparent, bien que de manière moins importante que dans la construction *dus4::ble*<sup>R</sup>. Cependant, il reste évident qu'il y a toujours bien plus de « sporulation defective cells » dans cette construction (25% de sporulation defective cells pour 45% d'asques correctement formés) comparativement à la souche wild type contrôle aussi comptée à ce moment-là (2% de sporulation defective cells), ce qui appuie le fait qu'introduire une cassette de résistance à ce locus provoque le défaut de sporulation.

Afin d'ensuite comprendre comment l'introduction de la cassette de résistance *ble*<sup>R</sup> peut perturber l'environnement génomique dans la construction  $\Delta dus4$  (*dus4::ble*<sup>R</sup>), un Northern Blot ciblant le gène *prp4* a été réalisé. *prp4* est un gène voisin proche de *dus4* situé sur le même brin d'ADN et séparé uniquement de 137 nucléotides (données provenant de [www.pombase.org](http://www.pombase.org)). Il a particulièrement retenu l'attention car son expression apparaissait fortement diminuée lors du RNAseq préalablement effectué sur la souche  $\Delta dus4$  (données non montrées). De plus, ce gène code pour une kinase essentielle (Prp4) qui joue un rôle dans les phénomènes d'épissage d'introns non-canoniques<sup>48</sup>, et dans le phénomène de ségrégation des chromosomes au cours de la méiose<sup>47</sup>. Prp4 est impliquée au cours de la première division méiotique. Un premier constat peut être fait : la présence du gène *prp4* (1900 nt) dans toutes les souches : WT,  $\Delta dus4$ ,  $\Delta dus4$  (leu-),  $\Delta dus2$ ,  $\Delta dus1-2-3-dus4$ <sub>DEAD</sub>, *dus4*<sub>STOP</sub>,  $\Delta dus1-2-3-4$ . Etant donné que le Northern Blot est brin non spécifique, c'est peut-être le long ARN non codant *SPNCRNA.1212.1*, qui fait face à *prp4* sur le brin anti-sens, qui est ici ciblé. Mais, étant donné que *SPNCRNA.1212.1* n'apparaît pas dans une souche  $\Delta dus4$  dans le Northern Blot *dus4* brin non spécifique (mais bien dans les autres souches), et qu'une bande est bien présente dans la souche  $\Delta dus4$  dans le Northern Blot *prp4*, il est très probable que la bande située à 1900 nt corresponde bien à *prp4* et non à *SPNCRNA.1212.1*. De plus, en se fiant aux données du Northern Blot *dus4*, le long ARN non codant possède une taille plus petite, d'environ 1700 nt. Les bandes situées sous *prp4* dans le Northern Blot *prp4* pourraient donc correspondre à *SPNCRNA.1212.1*.



Pour expliquer le fait que plusieurs bandes de différentes tailles soient marquées, il peut être de nouveau émis comme hypothèse qu'il existe non pas un ARN non codant comme référencé sur le site [www.pombase.org](http://www.pombase.org) (de 2905nt), mais bien plusieurs ARN non codants. En effet, l'annotation erronée d'ARN non codants a déjà été observée au laboratoire. Les longs ARN non codants jouent un rôle dans la régulation de l'expression des gènes et sont importants dans beaucoup de phénomènes physiologiques, mais aussi pathologiques (ex : cancer)<sup>53</sup>. Concernant les ARNm, les ARNlnc peuvent plus précisément impacter leur épissage, leur transport, leur traduction en protéine et leur dégradation. Ils peuvent jouer leur rôle en *cis* ou en *trans*<sup>53</sup>. Cependant, seulement quelques ARNlnc sont déjà identifiés et caractérisés chez *S. pombe*.

Deux autres constats peuvent être fait en analysant le Northern Blot *prp4*. Premièrement, contrairement à ce qui est ressorti au RNAseq, il n'est pas possible ici de discerner une diminution d'expression de *prp4* dans les souches délétées en *dus4*. Deuxièmement, dans toutes ces souches, une bande supplémentaire apparaît, de taille supérieure à *prp4*. En venant introduire une cassette de résistance au niveau de *dus4*, un ARN supplémentaire pouvant être associé au locus *prp4* est donc généré. L'hypothèse la plus probable serait qu'en venant introduire la cassette de résistance *ble<sup>R</sup>*, des ARN anti-sens correspondant à des versions plus ou moins tronquées ou modifiées de *prp4* seraient générés, à cause du promoteur du gène *ble<sup>R</sup>*.

Il est connu que des ARN anti-sens (asRNA) peuvent agir comme éléments régulateurs de l'expression de l'ARN, notamment en agissant de concert avec des facteurs en *trans*<sup>54</sup>. Les mécanismes post-transcriptionnels de ces interactions ARN-ARN peuvent donc jouer un rôle central dans la régulation de l'expression génique<sup>55</sup>. De plus, une étude sortie en 2011 de *Bitton et al*<sup>56</sup> a démontré qu'une vague de transcrits anti-sens non codants était générée lors de la différenciation sexuelle (notamment l'anti-sens de *spo6*). Ici, *prp4* étant un gène important pour la méiose, ce phénomène serait perturbé, et cela pourrait passer par une perturbation de l'expression du gène *spo6*. Des sporulation defective cells seraient ainsi observées. Les bandes de plusieurs tailles pourraient elles aussi correspondre à des ARN anti-sens de différentes tailles générées, ou, pour rappel, à des versions multiples du long ARN non codant *SPNCRNA.1212.1*. Une manière de vérifier cette hypothèse serait de réaliser un Northern Blot *prp4* brin spécifique, chose qui n'a pas été réalisée au cours de ce mémoire. De plus, l'étude de *Bitton et al*<sup>56</sup> rappelle aussi l'importance de considérer l'impact possible de la délétion d'un gène par l'introduction d'une cassette de résistance au niveau du gène voisin, et ce via une la provocation de la transcription du locus adjacent. Il est donc bien dangereux de prendre en compte un gène comme étant une identité génique indépendante à part entière quand on vient possiblement perturber la régulation de son environnement génique par l'introduction d'une cassette de résistance.

Une deuxième observation préalable à la réalisation de ce mémoire était la restauration du phénotype wild type dans la souche  $\Delta dus4 \Delta dus2$  ainsi que la restauration de l'expression du gène *spo6*. Dans ce mémoire, ce phénotype n'a pas été investigué de manière aussi poussée que le défaut de sporulation observé dans la souche  $\Delta dus4$ . Néanmoins, deux points ont été étudiés : la régulation d'expression compensatrice entre les gènes des enzymes *Dus2* et *Dus4*, et la possible nécessité de l'activité catalytique de *Dus2* lors de la restauration du phénotype wild type dans la souche  $\Delta dus4 \Delta dus2$ .

Concernant l'implication de l'activité catalytique de *Dus2*, une construction  $\Delta dus4 dus2_{DEAD}$  a été réalisée. Par comptages cellulaires, la souche  $\Delta dus4 dus2_{DEAD}$  sporule de façon similaire à une souche WT.



L'absence de l'activité catalytique de Dus2 est donc suffisante pour supprimer le phénotype du mutant de délétion de *dus4*. De manière surprenante, la suppression du défaut de sporulation qui est donc *dus4*-indépendant serait donc ici *dus2*-dépendante. Ensuite, en quantifiant le niveau d'expression de l'ARNm de *dus4* dans une souche  $\Delta$ *dus2*, il a pu être vérifié qu'il n'existe pas de régulation compensatrice d'expression de *dus4*, vu qu'il n'y a pas plus d'ARNm de *dus4* dans une souche  $\Delta$ *dus2*. Cela a également été vérifié par Northern Blot. Il aurait aussi été judicieux de quantifier l'expression de *prp4* dans la souche  $\Delta$ *dus4*  $\Delta$ *dus2* lors du Northern Blot *prp4*, pour visualiser la présence ou non de la bande supplémentaire située au-dessus de *prp4*. L'hypothèse que l'inactivation de Dus2 empêche la formation du transcrit anti-sens perturbant le gène *prp4* et la sporulation reste donc émise.

Il s'avère donc, selon toutes ces observations, que le défaut de sporulation observé dans la souche  $\Delta$ *dus4* corrélé à une diminution d'expression du gène *spo6* n'est pas attribuable à la perte de la dihydrouridine formée par l'enzyme Dus4. Au contraire, la perturbation de l'expression du gène voisin *prp4* important dans la méiose est l'hypothèse la plus probable à ce stade. Afin de le confirmer, des Northern Blot brin spécifique ainsi que des quantifications qRT-PCR sont nécessaires.

#### V.2 Préparation d'une analyse épitranscriptomique comparative de la dihydrouridine chez *Schizosaccharomyces pombe*, *Escherichia coli* et *Caenorhabditis elegans*

La deuxième partie de ce mémoire concernant l'étude de la dihydrouridine avait pour but d'étendre la technique de « D-sequencing » créée par Olivier Finet au cours de sa thèse, sur deux organismes modèles encore non « D-séquencés » : *Escherichia coli* et *Caenorhabditis elegans*. Cette technique vise à séquencer tout le transcriptome et à y déterminer les différentes positions de dihydrouridine sur l'ARN. Faisant suite à la cartographie de la position des dihydrouridine sur le transcriptome de *Schizosaccharomyces pombe*, les résultats futurs permettront une comparaison phylogénétique, entre un organisme eucaryote unicellulaire, un organisme procaryote et un organisme eucaryote pluricellulaire.

Olivier Finet ayant déjà mis au point cette technique sur *S. pombe*, il restait à transposer à deux autres organismes : *E. coli* et *C. elegans*, avant de pouvoir concrètement lancer le D-sequencing en tant que tel. La difficulté ici résidait dans le fait que la biologie de ces organismes est différente. L'extraction de l'ARN et son marquage par la rhodamine en vue de bloquer la rétro-transcription lors du D-seq n'étaient pas garantie d'autant de succès que pour *S. pombe*, en utilisant la même méthode. En effet, les transcrits d'*E. coli* sont généralement considérés comme moins stables vu que directement traductibles (*E. coli* ne possède pas de noyau). Il pouvait donc être difficile de les traiter avec le rhodamine labelling. Concernant *C. elegans*, vu la longueur importante des transcrits de cet organisme modèle, un bon labelling aurait pu être empêché par la présence de structures secondaires complexes. Afin de pouvoir lancer le rhodamine labelling, 30 µg d'ARN étaient nécessaires. Il fallait donc au moins récolter autant d'ARN lors de l'extraction d'ARN d'*E. coli* et de *C. elegans*. Afin d'obtenir des résultats les plus comparables possibles, le même protocole d'extraction d'ARN utilisé pour le D-seq de *S. pombe* a été testé. Au départ de 100 mL de LB à DO<sub>595nm</sub> 0,6-0,8 d'*E. coli* et d'environ 200 000 vers *C. elegans*, les quantités d'ARN récoltées se sont avérées suffisantes, validant ainsi la méthode d'extraction d'ARN utilisée et la quantité initiale d'*E. coli* et *C. elegans*. De plus, le profil d'analyse ARN obtenu par Bioanalyzer valide également la qualité de ces ARN obtenus.



Ensuite, du rhodamine labelling a été réalisé pour marquer la dihydrouridine. Après réalisation et quantification, la quantité d'ARN rhodaminisée s'est avérée suffisante et la variabilité par rapport à la quantité d'ARN de départ avant marquage acceptable vu le range de variabilité similaire obtenu précédemment chez *S. pombe* par Olivier Finet (données non montrées). Afin de vérifier et visualiser le marquage effectif de la dihydrouridine par la rhodamine, la technique de « Dot Blot assay » a été réalisée. Cette technique tire profit du fait que la rhodamine est un fluorophore. Les fluorescences émises par une goutte de chaque échantillon rhodaminisé (R+) et non rhodaminisé (échantillon contrôle, R-) ont été mesurées. Pour rappel, de la rhodamine est bien incorporée dans les échantillons R- et c'est l'absence de réduction au  $\text{NaBH}_4$  qui rend le marquage impossible. Concernant le dot blot de *C. elegans*, plusieurs constats peuvent être faits. Premièrement, la présence effective de plus de rhodamine dans les échantillons rhodaminisés comparativement à ceux qui ne le sont pas. Deuxièmement, le pourcentage de fluorescence émise par les échantillons R+ qui est d'environ 5% par les échantillons wild type rhodaminisés. Ce pourcentage est plus faible que lors des dot blot assay déjà réalisés précédemment par Olivier Finet sur *S. pombe* (en moyenne 14%, données non montrées). Même si la valeur d'expression est plus faible, il reste indiscutable que de la rhodamine est bien incorporée par rapport aux échantillons R-, et ce 5 fois plus en moyenne. Cela valide la technique de rhodamine labelling réalisée chez *C. elegans*. Cependant, ne disposant pas de souche contrôle délétée pour les dihydrouridine synthases, il n'est pas possible ici de contrôler le background de rhodamine incorporée par une autre modification que de la dihydrouridine. Concernant le dot blot d'*E. coli* plusieurs points sont également à discuter. Premièrement, comme pour *C. elegans*, la présence indiscutable de plus de rhodamine dans les échantillons R+ comparativement aux échantillons contrôles R- (60 fois plus dans la souche wild type et 27 fois plus dans les échantillons dépourvus de dihydrouridine synthase). Deuxièmement, le fait qu'il semble y avoir beaucoup plus de rhodamine incorporée dans les échantillons R+ de *E. coli* comparativement aux échantillons R+ de *C. elegans* et de *S. pombe*. Cela pourrait être expliqué par une présence plus importante de dihydrouridine chez *E. coli* comparativement à *C. elegans* et *S. pombe*. Il sera possible de le vérifier après obtention et analyse des résultats du D-seq. Un autre constat peut être fait : la présence plus importante de rhodamine incorporée dans les échantillons wild type comparativement aux échantillons délétés en dihydrouridine synthase (3 fois plus en moyenne). Chez *S. pombe*, cette différence était d'un peu moins de trois fois (données non montrées). Etant donné que la souche  $\Delta\text{dusABC}$  utilisée est clairement dépourvue de dihydrouridine synthase (pour rappel, souche provenant de l'étude de Bishop et al <sup>39</sup>), le background restant peut être attribué à autre chose qui bloquerait la rhodamine que la dihydrouridine (à moins qu'il n'y existe une dihydrouridine synthase supplémentaire). Il est connu dans la littérature que d'autres modifications de l'ARN sont sensibles au traitement utilisé lors du rhodamine labelling. D'autres modifications, comme la wybutosine (yW) et la 7-méthylguanosine ( $\text{m}^7\text{G}$ ) sont référencées comme pouvant incorporer des fluorophores après traitement au  $\text{NaBH}_4$  <sup>57</sup>. Cela pourrait partiellement expliquer le background restant de rhodamine dans les échantillons  $\Delta\text{dusABC}$ . A noter également l'aspect différent de la goutte contrôle de rhodamine dans le dot blot assay de *C. elegans* comparativement au dot blot assay d'*E. coli*. Cette différence d'aspect avait déjà aussi été observée lors de précédents dot blot assay réalisés par Olivier Finet sur *S. pombe*. Cela ne perturbe cependant pas l'interprétation des résultats ni la conclusion concernant la bonne efficacité du rhodamine labelling réalisé sur l'ARN extrait d'*E. coli* et de *C. elegans*.

Au terme de ce mémoire, il peut donc être confirmé que les protocoles d'extraction d'ARN et de rhodamine labelling réalisés chez *S. pombe* peuvent aussi être réalisés chez *E. coli* et *C. elegans*, en vue du D-seq. Il restera à réaliser et analyser le séquençage sur ces ARN.



A noter que tout comme dans l'étude de *Schwartz et al*<sup>34</sup>, l'efficacité de la méthode pourra être critiquée. Et ce en fonction des valeurs de D-ratio et de D-fold change obtenues, des quantifications de dihydrouridine et d'uridine dans les spike in insérés et ensuite quantifiés dans chaque librairie, des échantillons contrôles non rhodaminisés, de la souche contrôle  $\Delta dusABC$  dépourvue de dihydrouridine synthase dans le cas d'*E. coli* et du coverage du génome séquencé. Il faut aussi souligner le fait que dans le cas de *C. elegans*, c'est l'ARN de tous les tissus confondus de ce vers qui sera analysé. De plus, le coverage n'atteindra pas l'unité et aucune souche dépourvue de dihydrouridine synthase n'étant disponible, il manquera un contrôle important. En effet, la souche  $\Delta dusABC$  permet lors des analyses de résultats du D-seq d'éliminer les faux positifs potentiellement obtenus dans la souche wild type séquencée.



## VI. Conclusion et perspectives



## VI. Conclusion et perspectives

L'épitranscriptomique est un champ d'étude qui n'en est encore qu'à ses débuts prometteurs dans le monde de la recherche en génétique moléculaire. Au cours de ce mémoire, une modification post-transcriptionnelle de l'ARN encore peu étudiée jusqu'à présent a été étudiée chez la levure de fission *Schizosaccharomyces pombe* : la dihydrouridine.

Un des buts de ce mémoire de recherche était d'investiguer le rôle biologique de la dihydrouridine dans le phénomène de la sporulation de *S. pombe*, suite à l'observation d'un défaut de sporulation dans une souche dépourvue de dihydrouridine synthase 4. Ce phénotype était corrélé à un défaut d'expression du gène de sporulation *spo6*. Après analyse de différents mutants de *Dus4*, il est apparu que le défaut de sporulation n'est pas attribuable à l'absence de la protéine *Dus4*. L'hypothèse que nous retenons est la perturbation de l'expression d'un ou plusieurs ARN non codant quand le gène *dus4* est remplacé par une cassette de résistance à un antibiotique. Il en résulterait une dérégulation du gène voisin, *prp4*, impliqué à la fois dans l'épissage d'ARNm exprimés en méiose (dont potentiellement *spo6*) et à la fois dans la ségrégation des chromosomes. Indirectement, notre travail a donc peut-être révélé une régulation inconnue de l'expression de *prp4*.

Nous avons aussi établi que la délétion du gène *dus2*, ou l'inactivation de l'activité catalytique de l'enzyme *Dus2* suffisait à supprimer le phénotype du mutant  $\Delta$ *dus4*. Nous n'avons pas à ce jour d'hypothèse solide pour expliquer ces données.

Ce travail souligne l'importance d'avérer des données initiales interpellantes avant de poursuivre leur analyse et de bâtir des hypothèses. La conclusion ferme de la première partie de notre travail qui contredit l'hypothèse initiale est que l'enzyme *Dus4* n'est pas requise pour la sporulation. Les découvertes récentes des rôles d'ARN dans la régulation de l'expression génique (RNAi, lncRNA, asRNA), sont en accord avec le fait que perturber un gène en le remplaçant par une cassette de résistance puisse également perturber une régulation transcriptionnelle fine du locus.

Un autre aspect de l'épitranscriptomique, outre l'investigation des rôles biologiques des modifications post-transcriptionnelles, est la mise au point de techniques permettant de séquencer ces modifications. Jusqu'à présent, il n'existe pas de technique permettant de séquencer simultanément plusieurs modifications. Chez *S. pombe*, une technique de « D-sequencing » permettant de séquencer les différentes positions de la dihydrouridine au sein du transcriptome a été développée au laboratoire (Olivier Finet).

Le deuxième but de notre travail était ici de préparer le « D-Seq » de la dihydrouridine chez deux autres espèces modèles : *E. coli* et *C. elegans*. Les mises au point techniques ont confirmé que l'ARN de ces deux espèces modèles pouvait être extrait en suffisance et marqué par de la rhodamine au niveau de l'uridine de la même manière que l'ARN de *S. pombe*. L'envoi des ARN finaux extraits, rhodaminisés et préparés pour le D-seq sera réalisé par Olivier Finet lors de la suite de sa thèse sur l'étude de la dihydrouridine. A noter qu'il serait intéressant d'également disposer d'une souche contrôle de *C. elegans* dépourvue de dihydrouridine synthase, pour intégrer un contrôle R- qui garantira la spécificité de l'approche.



## VII. Annexes



## VII. Annexes

Table 1 : Souches de *Schizosaccharomyces Pombe* utilisées dans ce mémoire

Souche	Désignation	Génotype	Source
568	WT	h <sup>90</sup>	P. Nurse
973	<i>ura4-D18</i>	h <sup>90</sup> <i>ura4-D18</i>	S. Fauquenoy
1196	<i>hst2::ble<sup>R</sup></i>	h <sup>90</sup> <i>hst2::ble<sup>R</sup></i>	S. Fauquenoy
1445	$\Delta$ <i>dus2</i>	h <sup>90</sup> <i>dus2::kan<sup>R</sup></i>	Bioneer
1494	$\Delta$ <i>dus1-2-3-4</i>	h <sup>90</sup> <i>dus1::nat<sup>R</sup> dus2::kan<sup>R</sup> dus3::hph<sup>R</sup> dus4::ble<sup>R</sup></i>	V. Migeot
1544	$\Delta$ <i>dus2-4</i>	h <sup>90</sup> <i>dus2::kan<sup>R</sup> dus4::ble<sup>R</sup></i>	V.Migeot
1631	$\Delta$ <i>dus4</i>	h <sup>90</sup> <i>dus4::ble<sup>R</sup></i>	C. Yague-Sanz
1795	$\Delta$ <i>dus1-3-4</i>	h <sup>90</sup> <i>dus1::nat<sup>R</sup> dus3::hph<sup>R</sup> dus4::ble<sup>R</sup></i>	O.Finet
1818	<i>dus4::ura4 ura4-D18</i>	h <sup>90</sup> <i>dus4::ura4 ura4-D18</i>	O. Finet
1819	<i>dus4<sub>DEAD</sub> ura4-D18</i>	h <sup>90</sup> <i>dus4 C108A K149A ura4-D18</i>	O. Finet
1820	<i>dus4<sub>DEAD</sub></i>	h <sup>90</sup> <i>dus4 C108A K149A</i>	O.Finet
1843	$\Delta$ <i>dus1-2-3 dus4<sub>DEAD</sub></i>	h <sup>90</sup> <i>dus1::nat<sup>R</sup> dus2::kan<sup>R</sup> dus3::hph<sup>R</sup> dus4 C108A K149A</i>	Mémoire
1845	<i>dus4<sub>DEAD</sub>-TAP</i>	h <sup>90</sup> <i>dus4 C108A K149A-TAP-nat<sup>R</sup></i>	Mémoire
1846	<i>dus4<sub>STOP</sub> ura4-D18</i>	h <sup>90</sup> <i>dus4<sub>STOP</sub> A28T G37T A46T C92G T158A G499T G584T T707A ura4-D18</i>	Mémoire
1847	<i>dus4-TAP</i>	h <sup>90</sup> <i>dus4-TAP-nat<sup>R</sup></i>	Mémoire
1848	<i>dus2::ura4 ura4 D-18</i>	h <sup>90</sup> <i>dus2::ura4 ura4-D18</i>	Mémoire
1852	<i>dus4<sub>STOP</sub></i>	h <sup>90</sup> <i>dus4<sub>STOP</sub> A28T G37T A46T C92G T158A G499T G584T T707A</i>	Mémoire
1853	<i>dus2<sub>DEAD</sub> ura4-D18</i>	h <sup>90</sup> <i>dus2 C108A K149A ura4-D18</i>	Mémoire
1854	<i>dus4::hph<sup>R</sup></i>	h <sup>90</sup> <i>dus4::hph<sup>R</sup></i>	Mémoire
1858	<i>dus4<sub>STOP</sub>-TAP</i>	h <sup>90</sup> <i>dus4<sub>STOP</sub> A28T G37T A46T C92G T158A G499T G584T T707A -TAP-kan<sup>R</sup></i>	Mémoire
1868	<i>dus2<sub>DEAD</sub></i>	h <sup>90</sup> <i>dus2 C108A K149A</i>	Mémoire
1899	$\Delta$ <i>dus4 dus2<sub>DEAD</sub></i>	h <sup>90</sup> <i>dus2 C116A K159A dus4::ble<sup>R</sup></i>	Mémoire
OF12	$\Delta$ <i>dus4<sub>BIS</sub></i>	h <sup>90</sup> <i>dus4::ble<sup>R</sup></i>	Mémoire



Table 2 : Primers utilisés dans ce mémoire

Construction	Technique	Forward Primer (5'-3')	Reverse Primer (5'-3')
<i>Adus4<sup>BIS</sup></i>	High-fidelity PCR	ACTGCGTAGATGTCACATATG CTGTCAATGGCCGTGGACATT TTAGCAATTGCCAACATT- <b>CGGATCCCCGGGTAAATTAA</b>	TGATTACAGAAATCACTTCATAT TTACAATGCCAACTAATAAACAA ACAATGAGTTGAGA- <b>GAATTCGAGCTCGTTTAAAC</b>
	PCR colony	TAGTAGAAACGCCAACTGG	CGGATGTGATGTGAGAACTGTAT CCTAGC
<i>dus4::ura4</i>	High-fidelity PCR	AGAGAAGCTGGTTGGAAGGC	GAGAGATCGTTTAAAGGATCC
<i>dus4<sup>DEAD</sup></i>	High-fidelity PCR	CAGTAATTTTGAAAAGGCGG CACAACCTGTTGGACCCTATG TAGATGGTATTGGA- <b>CGCCAGGGTTTCCCAGTCA CGAC</b>	CTCCAGATTTTCAATGACTTGC ATCAGGTGCCGAGTTTATTTAA GTCTTTGGC- <b>AGCGGATAACAATTTACACAG GA</b>
	PCR colony <sub>1</sub>	AGAGAAGCTGGTTGGAAGC	GAGAGATCGTTTAAAGGATCC
	gBLOCK amplification	GAGAGATCGTTTAAAGGATC	GATATGAAGAAACGCCTTGG
	PCR colony <sub>2</sub>	TAGTAGAAACGCCAACTGG	CCAAGGCGTTTCTTCATATC
<i>dus2<sup>DEAD</sup></i>	High-fidelity PCR	GCTAATCGATTGATTTTCAA CTCGGAAGTGCATCCCCAGAA CTTGCACTGGAAGCAGCA- <b>CGCCAGGGTTTCCCAGTCA CGAC</b>	TGTATCGCATATGCGCTCTACCA ATTTTAGGGTATCTTCTTTGTCT CTAATAAACGAAT- <b>AGCGGATAACAATTTACACAG GA</b>
	PCR colony <sub>1</sub>	CAGTAATTTTGAAAAGGCGG CACAACCTGTTGGACCCTATG TAGATGGTATTGGA- <b>CGCCAGGGTTTCCCAGTCA CGAC</b>	TATTTTCGACAGACGCCTAC
	gBLOCK amplification	TCAGTTATTATTGACTCAATTT AAAATAG	TATTTTCGACAGACGCCTAC
	PCR colony <sub>2</sub>	GCAAAGATGACTTTTGAGTG	TATTTTCGACAGACGCCTAC
<i>Adus1-2-3- dus4<sup>DEAD</sup></i>	PCR colony	<i>Dus1-2-3</i> : CGGATGTGATGTGAGAACTGT ATCCTAGC	<i>dus1</i> :GCTACAAATCTACAAAATCC TGCC <i>dus2</i> :GAGAATCGTTAATTGTGTCC G <i>dus3</i> :TGTGTACCTAAAAGCATATA GCG
	PCR colony <sub>1</sub>	TAGTAGAAACGCCAACTGG	CGGATGTGATGTGAGAACTGTAT CCTAGC
	PCR colony <sub>2</sub>	AGGTTTATCTTTTACTTCTAG AGG	GGACAAAGGATATATTGATTTTA CG
<i>dus4<sup>STOP</sup></i>	gBLOCK amplification	CGAAATAAAACCAATGGTAA CG	AGTCATTCTTTCACGCTTAG
	PCR colony <sub>1</sub>	TAGTAGAAACGCCAACTGG	TTAATAGCGGCGACAAGGC
	PCR colony <sub>2</sub>	GAGAGATCGTTTAAAGGATCC	CCAAGGCGTTTCTTCATATC
<i>dus4<sup>STOP</sup>-TAP, dus4<sup>DEAD</sup>TAP, dus4-TAP</i>	High-fidelity PCR	GTTGTTTCGGAGACCCGATGAA CCCATGTTTGGAGAATCGGTA TTGCCTTGTCGCCGCTAT - <b>CGGATCCCCGGGTAAATTAA</b>	TGATTACAGAAATCACTTCATAT TTACAATGCCAACTAATAAACAA ACAATGAGTTGAGA - <b>GAATTCGAGCTCGTTTAAAC</b>
	PCR colony	GTTCGCCCATGTGTCCAAAT	CAAGGAGGGTATTCTGGGCC
<i>dus4::hph<sup>R</sup></i>	High-fidelity PCR	ACTGCGTAGATGTCACATATG CTGTCAATGGCCGTGGACATT T TAGCAATTGCCAACATT- <b>CGGATCCCCGGGTAAATTAA</b>	TGATTACAGAAATCAC TTCATATTTACAATGC CAACTAATAAACAAACA ATGAGTTGAGA-GAATTC <b>GAGCTCGTTTAAAC</b>



Table 2 : Primers utilisés dans ce mémoire : suite

Construction	Technique	Forward Primer (5'-3')	Reverse Primer (5'-3')
<i>dus2</i> <sub>DEAD</sub> <i>Δdus4</i>	PCR colony <sub>1</sub>	GCAAAGATGACTTTTGAGTG	TATTTCGACAGACGCCTAC
	PCR colony <sub>2</sub>	TAGTAGAAACGCCAACTGG	CGGATGTGATGTGAGAACTGTATCCTAGC
<i>spo6</i>	Northern Blot probe	AAC TATTGAACAGAATCCTC	CCTTTCCAAGTCCTTGTATC
	qRT-PCR	CTGGTAGAAGAAGGGAGTTA	AACAACGTTATTTTGGTTAG
<i>dus4</i>	Northern Blot probe	GAGAGATCGTTTAAAGGATCC	CCAAGGCGTTTCTTCATATC
	qRT-PCR	CGAAATAAAACCAATGGTAACG	AGTCATTCTTTCACGCTTAG
<i>prp4</i>	Northern Blot probe	ATCAATTCTTGAGCGACTGC	GTATCTTCAATTTGTACAGAAGA
<i>act1</i>	qRT-PCR	CCACTATGTATCCCGGTATTGC	CAATCTTGACCTTCATGGAGCT

*Remarque* : les séquences s'hybridant sur un vecteur sont surlignées en gras



## VIII. Bibliographie



## VIII. Bibliographie

- 1 Davis, F. F. & Allen, F. W. Ribonucleic acids from yeast which contain a fifth nucleotide. *J Biol Chem* **227**, 907-915 (1957).
- 2 Song, J. & Yi, C. Chemical Modifications to RNA: A New Layer of Gene Expression Regulation. *ACS Chem Biol* **12**, 316-325, doi:10.1021/acscchembio.6b00960 (2017).
- 3 Alfonzo, J. D. Post-transcriptional modifications are very important after all. *RNA Biol* **11**, 1481-1482, doi:10.1080/15476286.2014.1005439 (2014).
- 4 Laptev, I. G., Golovina, A. Y., Sergiev, P. V. & Dontsova, O. A. [Posttranscriptional messenger RNA modifications in eukaryotes]. *Mol Biol (Mosk)* **49**, 923-936, doi:10.7868/S0026898415060142 (2015).
- 5 Machnicka, M. A. *et al.* MODOMICS: a database of RNA modification pathways--2013 update. *Nucleic Acids Res* **41**, D262-267, doi:10.1093/nar/gks1007 (2013).
- 6 Zhao, B. S., Roundtree, I. A. & He, C. Post-transcriptional gene regulation by mRNA modifications. *Nat Rev Mol Cell Biol* **18**, 31-42, doi:10.1038/nrm.2016.132 (2017).
- 7 Frye, M., Jaffrey, S. R., Pan, T., Rechavi, G. & Suzuki, T. RNA modifications: what have we learned and where are we headed? *Nat Rev Genet* **17**, 365-372, doi:10.1038/nrg.2016.47 (2016).
- 8 Harcourt, E. M., Kietrys, A. M. & Kool, E. T. Chemical and structural effects of base modifications in messenger RNA. *Nature* **541**, 339-346, doi:10.1038/nature21351 (2017).
- 9 Lewis, C. J., Pan, T. & Kalsotra, A. RNA modifications and structures cooperate to guide RNA-protein interactions. *Nat Rev Mol Cell Biol* **18**, 202-210, doi:10.1038/nrm.2016.163 (2017).
- 10 Sprinzl, M., Horn, C., Brown, M., Ioudovitch, A. & Steinberg, S. Compilation of tRNA sequences and sequences of tRNA genes. *Nucleic Acids Res* **26**, 148-153 (1998).
- 11 Byrne, R. T. *et al.* Major reorientation of tRNA substrates defines specificity of dihydrouridine synthases. *Proc Natl Acad Sci U S A* **112**, 6033-6037, doi:10.1073/pnas.1500161112 (2015).
- 12 Cantara, W. A. *et al.* The RNA Modification Database, RNAMDB: 2011 update. *Nucleic Acids Res* **39**, D195-201, doi:10.1093/nar/gkq1028 (2011).
- 13 Whelan, F. *et al.* From bacterial to human dihydrouridine synthase: automated structure determination. *Acta Crystallogr D Biol Crystallogr* **71**, 1564-1571, doi:10.1107/S1399004715009220 (2015).
- 14 Dalluge, J. J., Hashizume, T., Sopchik, A. E., McCloskey, J. A. & Davis, D. R. Conformational flexibility in RNA: the role of dihydrouridine. *Nucleic Acids Res* **24**, 1073-1079 (1996).
- 15 Xing, F., Hiley, S. L., Hughes, T. R. & Phizicky, E. M. The specificities of four yeast dihydrouridine synthases for cytoplasmic tRNAs. *J Biol Chem* **279**, 17850-17860, doi:10.1074/jbc.M401221200 (2004).
- 16 Kasprzak, J. M., Czerwonec, A. & Bujnicki, J. M. Molecular evolution of dihydrouridine synthases. *BMC Bioinformatics* **13**, 153, doi:10.1186/1471-2105-13-153 (2012).
- 17 Dalluge, J. J. *et al.* Posttranscriptional modification of tRNA in psychrophilic bacteria. *J Bacteriol* **179**, 1918-1923 (1997).
- 18 Xing, F., Martzen, M. R. & Phizicky, E. M. A conserved family of *Saccharomyces cerevisiae* synthases effects dihydrouridine modification of tRNA. *RNA* **8**, 370-381 (2002).



- 19 Yu, F. *et al.* Molecular basis of dihydrouridine formation on tRNA. *Proc Natl Acad Sci U S A* **108**, 19593-19598, doi:10.1073/pnas.1112352108 (2011).
- 20 Savage, D. F., de Crecy-Lagard, V. & Bishop, A. C. Molecular determinants of dihydrouridine synthase activity. *FEBS Lett* **580**, 5198-5202, doi:10.1016/j.febslet.2006.08.062 (2006).
- 21 Park, F. *et al.* Crystal structure of YIGZ, a conserved hypothetical protein from *Escherichia coli* k12 with a novel fold. *Proteins* **55**, 775-777, doi:10.1002/prot.20087 (2004).
- 22 Kato, T. *et al.* A novel human tRNA-dihydrouridine synthase involved in pulmonary carcinogenesis. *Cancer Res* **65**, 5638-5646, doi:10.1158/0008-5472.CAN-05-0600 (2005).
- 23 Nurse, P. Genetic control of cell size at cell division in yeast. *Nature* **256**, 547-551 (1975).
- 24 Hoffman, C. S., Wood, V. & Fantes, P. A. An Ancient Yeast for Young Geneticists: A Primer on the *Schizosaccharomyces pombe* Model System. *Genetics* **201**, 403-423, doi:10.1534/genetics.115.181503 (2015).
- 25 Hagan, I. M., Grallert, A. & Simanis, V. Analysis of the *Schizosaccharomyces pombe* Cell Cycle. *Cold Spring Harb Protoc* **2016**, pdb top082800, doi:10.1101/pdb.top082800 (2016).
- 26 Yamashita, A., Sakuno, T., Watanabe, Y. & Yamamoto, M. Analysis of *Schizosaccharomyces pombe* Meiosis. *Cold Spring Harb Protoc* **2017**, pdb top079855, doi:10.1101/pdb.top079855 (2017).
- 27 Otsubo, Y. & Yamamoto, M. Signaling pathways for fission yeast sexual differentiation at a glance. *J Cell Sci* **125**, 2789-2793, doi:10.1242/jcs.094771 (2012).
- 28 Nakamura, T., Kishida, M. & Shimoda, C. The *Schizosaccharomyces pombe* spo6+ gene encoding a nuclear protein with sequence similarity to budding yeast Dbf4 is required for meiotic second division and sporulation. *Genes Cells* **5**, 463-479 (2000).
- 29 Nakamura, T., Nakamura-Kubo, M., Nakamura, T. & Shimoda, C. Novel fission yeast Cdc7-Dbf4-like kinase complex required for the initiation and progression of meiotic second division. *Mol Cell Biol* **22**, 309-320 (2002).
- 30 Li, X., Xiong, X. & Yi, C. Epitranscriptome sequencing technologies: decoding RNA modifications. *Nat Methods* **14**, 23-31, doi:10.1038/nmeth.4110 (2016).
- 31 Dominissini, D. *et al.* Topology of the human and mouse m6A RNA methylomes revealed by m6A-seq. *Nature* **485**, 201-206, doi:10.1038/nature11112 (2012).
- 32 Meyer, K. D. *et al.* Comprehensive analysis of mRNA methylation reveals enrichment in 3' UTRs and near stop codons. *Cell* **149**, 1635-1646, doi:10.1016/j.cell.2012.05.003 (2012).
- 33 Hsu, P. J., Shi, H. & He, C. Epitranscriptomic influences on development and disease. *Genome Biol* **18**, 197, doi:10.1186/s13059-017-1336-6 (2017).
- 34 Schwartz, S. *et al.* Transcriptome-wide mapping reveals widespread dynamic-regulated pseudouridylation of ncRNA and mRNA. *Cell* **159**, 148-162, doi:10.1016/j.cell.2014.08.028 (2014).
- 35 Novoa, E. M., Mason, C. E. & Mattick, J. S. Charting the unknown epitranscriptome. *Nat Rev Mol Cell Biol* **18**, 339-340, doi:10.1038/nrm.2017.49 (2017).
- 36 Motorin, Y., Muller, S., Behm-Ansmant, I. & Branlant, C. Identification of modified residues in RNAs by reverse transcription-based methods. *Methods Enzymol* **425**, 21-53, doi:10.1016/S0076-6879(07)25002-5 (2007).
- 37 Soltis, D. E. & Soltis, P. S. The role of phylogenetics in comparative genetics. *Plant Physiol* **132**, 1790-1800 (2003).



- 38 Ellegren, H. Comparative genomics and the study of evolution by natural selection. *Mol Ecol* **17**, 4586-4596, doi:10.1111/j.1365-294X.2008.03954.x (2008).
- 39 Bishop, A. C., Xu, J., Johnson, R. C., Schimmel, P. & de Crecy-Lagard, V. Identification of the tRNA-dihydrouridine synthase family. *J Biol Chem* **277**, 25090-25095, doi:10.1074/jbc.M203208200 (2002).
- 40 Squiban, B. & Kurz, C. L. C. elegans: an all in one model for antimicrobial drug discovery. *Curr Drug Targets* **12**, 967-977 (2011).
- 41 Hillier, L. W. *et al.* Genomics in C. elegans: so many genes, such a little worm. *Genome Res* **15**, 1651-1660, doi:10.1101/gr.3729105 (2005).
- 42 Brown, S. D. & Lorenz, A. Single-step Marker Switching in Schizosaccharomyces pombe Using a Lithium Acetate Transformation Protocol. *Bio Protoc* **6**, doi:10.21769/BioProtoc.2075 (2016).
- 43 Betteridge, T. *et al.* Fluorescent labeling of tRNAs for dynamics experiments. *RNA* **13**, 1594-1601, doi:10.1261/rna.475407 (2007).
- 44 Grimm, C., Kohli, J., Murray, J. & Maundrell, K. Genetic engineering of Schizosaccharomyces pombe: a system for gene disruption and replacement using the ura4 gene as a selectable marker. *Mol Gen Genet* **215**, 81-86 (1988).
- 45 Vogeli, G. The nucleotide sequence of tRNA tyrosine from the fission yeast Schizosaccharomyces pombe. *Nucleic Acids Res* **7**, 1059-1065 (1979).
- 46 Wierman, M. B. & Smith, J. S. Yeast siruins and the regulation of aging. *FEMS Yeast Res* **14**, 73-88, doi:10.1111/1567-1364.12115 (2014).
- 47 Pozgajova, M., Cipak, L. & Trakovicka, A. Prp4 kinase is required for proper segregation of chromosomes during meiosis in Schizosaccharomyces pombe. *Acta Biochim Pol* **60**, 871-873 (2013).
- 48 Eckert, D. *et al.* Prp4 Kinase Grants the License to Splice: Control of Weak Splice Sites during Spliceosome Activation. *PLoS Genet* **12**, e1005768, doi:10.1371/journal.pgen.1005768 (2016).
- 49 Llorca, O. *et al.* Eukaryotic chaperonin CCT stabilizes actin and tubulin folding intermediates in open quasi-native conformations. *EMBO J* **19**, 5971-5979, doi:10.1093/emboj/19.22.5971 (2000).
- 50 Kowalak, J. A., Bruenger, E. & McCloskey, J. A. Posttranscriptional modification of the central loop of domain V in Escherichia coli 23 S ribosomal RNA. *J Biol Chem* **270**, 17758-17764 (1995).
- 51 Sanger, F., Nicklen, S. & Coulson, A. R. DNA sequencing with chain-terminating inhibitors. *Proc Natl Acad Sci U S A* **74**, 5463-5467 (1977).
- 52 Siwaszek, A., Ukleja, M. & Dziembowski, A. Proteins involved in the degradation of cytoplasmic mRNA in the major eukaryotic model systems. *RNA Biol* **11**, 1122-1136, doi:10.4161/rna.34406 (2014).
- 53 Mathieu, E. L., Belhocine, M., Dao, L. T., Puthier, D. & Spicuglia, S. [Functions of lncRNA in development and diseases]. *Med Sci (Paris)* **30**, 790-796, doi:10.1051/medsci/20143008018 (2014).
- 54 Nishizawa, M., Ikeya, Y., Okumura, T. & Kimura, T. Post-transcriptional inducible gene regulation by natural antisense RNA. *Front Biosci (Landmark Ed)* **20**, 1-36 (2015).
- 55 Roignant, J. Y. & Soller, M. m6A in mRNA: An Ancient Mechanism for Fine-Tuning Gene Expression. *Trends Genet*, doi:10.1016/j.tig.2017.04.003 (2017).
- 56 Bitton, D. A. *et al.* Programmed fluctuations in sense/antisense transcript ratios drive sexual differentiation in S. pombe. *Mol Syst Biol* **7**, 559, doi:10.1038/msb.2011.90 (2011).



- 57 Behm-Ansmant, I., Helm, M. & Motorin, Y. Use of specific chemical reagents for detection of modified nucleotides in RNA. *J Nucleic Acids* **2011**, 408053, doi:10.4061/2011/408053 (2011).
- 58 Taj, Muhammad & Samreen, Zohra & Ling, Ji & Hassani, Imran & Hassani, Taj & Yunlin, Wei. (2014). ESCHERICHIA COLI AS A MODEL ORGANISM. *Int. J. Engg. Res. & Sci. & Tech.* 3. 1-8.